A NOVEL APPROACH TO THE RATIONAL DESIGN OF ARTIFICIAL ENZYMES

A Thesis presented by

Catherine Elizabeth Atkinson

In Partial Fulfilment of the Requirements
for the Award of the Degree of

DOCTOR OF PHILOSOPHY

OF THE

UNIVERSITY OF LONDON

Christopher Ingold Laboratories
Department of Chemistry
University College London
London WC1H 0AJ

December 2001

ABSTRACT

This thesis describes a new approach to the rational design of an artificial esterase, and is in three parts.

The first part of the thesis is an introduction to the design and synthesis of artificial enzymes. This details both traditional 'design', and more recent 'selection' approaches to artificial enzymes, and discusses the advantages and problems associated with the different strategies.

The second part is a discussion of our investigations into the design and synthesis of polymeric artificial enzymes which are able to catalyse an ester hydrolysis. The research uses a novel strategy that combines the 'design' and 'selection' approaches: Using the concept of transition state theory, we aimed to design a unit that should bind and stabilise the transition state of a reaction. We then aimed to selectively incorporate this, and other catalytically active groups into a flexible, soluble polymer. The work starts with the synthesis of the ester substrate, and a phosphonate transition state analogue.

The 'design' section of the project involved finding a low molecular weight 'binding unit' that could bind to the transition state analogue, and hence should bind and stabilise the transition state of our reaction. We decided to use dipeptides and studied binding using Pulsed Field Gradient NMR techniques to measure changes in diffusion coefficients. This is a technique which had previously only been used to study large molecule-small molecule binding. We successfully managed to apply this technique to studying small molecule-small molecule interactions, and thus it was found that the dipeptide arginine-arginine bound well to our transition state analogue, and the nature of the interactions were studied using molecular modelling.

The 'selection' section involved incorporation of this unit into a polymer, along with the introduction of amino acids that could act as catalytically active groups. Both the dipeptide, and chosen amino acids were reacted with a polyallylamine backbone using standard peptide chemistry. The influences of the different polymers on the hydrolysis of the ester were investigated, and it was found that some of these polymers showed distinct catalytic 'enzyme like' properties. The results are reported within.

The third section described the experimental work and procedure used throughout this thesis.

Table of Contents

Acknowledgements.

I would like to thank my supervisor Prof. Willie Motherwell for giving me the opportunity to work on this challenging project, and for his help, encouragement and belief in me over the past three years. Thanks also to Stephanie Wong and Brian Warrington from Smithkline Beecham for sharing their expertise with me.

My project has covered a broad range of subjects, and as a result there are several people without whose help and advice I would have been lost on several occasions. Two people in particular deserve special thanks: Dr. Abil Aliev for all the time, effort and care he put into setting up and helping run all the NMR studies, and Dr. Grant Wishart, from the Wolfson Institute who not only let me use his computers, but took time to teach me how to run the molecular modelling studies. I'd also like to thank Prof. Tony Barrett at Imperial College for letting me use his microscope IR, Dr. David Selwood and Dick Champion, from the Wolfson Institute for their advice on peptide synthesis, Laura Williams from Smithkline Beecham and Steve Corker at UCL for their help with HPLC. I also thank all the technical staff at UCL for their help: Jill Maxwell and Alan Stones for microanalysis, Jorge Gonzales for NMR, and John Hill for mass spectra. I owe a huge debt to my three proof readers, Beatrice, Oliver and Camilla, who have done a terrific job in a very short space of time. And of course thanks to both Smithkline Beecham and the EPSRC for funding.

In the lab I would like to give special thanks to Beatrice, for being the best bench mate I could have hoped for, and for her friendship, advice and incredible patience in listening to me moaning about polymers, and Tilly, for welcoming me into the lab, making my first year so enjoyable, and showing me that it is possible to make it through a PhD on artificial enzymes. Thanks to Lynda, Santi and Julien, three postdocs always willing to help, and to Robyn for keeping the lab in the good order it is. Other people who have made my time in the lab a pleasant experience include Camilla, for all the things she does and says that make me smile, Oliver for being a bottomless pit of information about everything, Rehan for willingly taking everything we threw at him, and Mike, Ray, Marta, Yvan, Rosa, Christoph, and anyone else I've forgotten along the way.

Outside the lab, special thanks to Claire, Phil, Pete and all "the boys" for always being there to take my mind off the lab when required.

Finally I'd like to thank Matthew and my parents for their continued love, support and encouragement. I couldn't have done this without them.

Abbreviations

Ac	acetyl
Ala	alanine
aq	aqueous
Ar	aryl
Arg	arginine
Asp	aspartic acid
b	broad
Boc	tertiary butoxycarbonyl
t-Bu	tertiary butyl
CA	6-aminocaproic acid
CD	cyclodextrin
CI	chemical ionisation
conc	concentrated
Cys	cysteine
d	doublet
DCC	*N, N'*-dicyclohexyl carbodiimide
DCL	dynamic combinatorial library
DCM	dichloromethane
DEPT	distortionless enhancement by polarisation transfer
DIC	*N, N'*-diisopropylcarbodiimide
DIPEA	diisopropylethylamine
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DOSY	Diffusion Ordered SpectroscopY
EDC	1-[3-(dimethylamino)propyl]-2-ethylcarbodiimide hydrochloride
ee	enantiomeric excess
EI	electron impact
equiv	molar equivalents
Et	ethyl
EtOAc	ethyl acetate

FAB	fast atom bombardment
Fmoc	9-fluorenylmethoxycarbonyl
Gly	glycine
HATU	*N*-[(dimethylamino)-1*H*-1,2,3-triazolo[4,5-*b*]pyridin-1-yl-methylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide
His	histidine
HOBt	[1-hydroxybenzoxytriazole]
HPLC	high performance liquid chromatography
IR	infra red
J	coupling constant
Leu	leucine
Lys	lysine
m	multiplet
m.p.	melting point
Me	methyl
Mins	minutes
MIP	molecularly imprinted polymer
NHS	*N*-hydroxysuccinimide
NMM	*N*-methylmorpholine
NMR	nuclear magnetic resonance
Pfb	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PFG	pulsed field gradient
Ph	phenyl
Phe	phenylalanine
ppm	parts per million
PyBOP	(1*H*-benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
q	quartet
Rf	retention factor
r.t.	room temperature
s	singlet
sat	saturated

Ser	serine
t	triplet
tert	tertiary
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
THF	tetrahydrofuran
Thr	threonine
TIS	triisopropylsilane
TLC	thin layer chromatography
TMS	trimethylsilyl
tren	tris(2-aminoethyl)amine
Trt	trityl
TSA	transition state analogue
TSP	2,2,3,3-d(4)-3(Trimethylsilyl)propionic acid sodium salt
Tyr	tyrosine
UV	ultra violet
w	weak
w/v	weight by volume

Chapter 1. Introduction.

Progress in the Design and Synthesis of Artificial Enzymes.

1. Introduction.

Nature has developed enzymes as highly evolved catalysts which not only accelerate reactions, but which can also do so with near-perfect regio and stereo specificity. As well as providing an intellectual challenge, a proper understanding of the underlying principles of enzyme catalysis would enable us to develop 'artificial enzymes' which would act with the same efficiency as enzymes, and have some considerable advantages over the natural systems. These include the potential to design catalysts for systems for which either no natural enzymes are known, or are available for synthetic application. As enzymes are proteins, they generally operate in aqueous media in order to retain the 3- dimensional structure required for catalysis. Artificial enzymes may not have this same requirement and therefore be better suited to use in organic solvents. Since the protein structure of enzymes is fragile, artificial enzymes can in principle be synthesised from macromolecular structures which are more chemically robust, and therefore could be utilised at higher temperatures and pressures.

Nature has of course had billions of years to evolve enzymes into the complex 3-dimensional structures with which we are familiar today. As synthetic chemists however, we need to concentrate this period into a feasible timescale for research. In order to be able to do this successfully, the first requirement is to have some understanding of the concepts that lie behind enzyme catalysis.

2. Determinant Factors in Enzymic Catalysis.

As early as 1894, Emil Fischer proposed that catalysis arose as a result of binding between the enzyme and substrate, a proposal he named the lock and key theory[1]. Indeed, it is now known that a characteristic feature of an enzyme catalysed reaction is that it follows Michealis-Menton kinetics due to initial formation of an enzyme substrate complex[2].

Today, the most commonly accepted view is that proposed by Pauling in 1948, who suggested that catalysis occurs due to the enzyme's ability to stabilise the transition-state structure of the substrate relative to that of the ground state[3]. The theory is illustrated below for a unimolecular reaction (Figure 1).

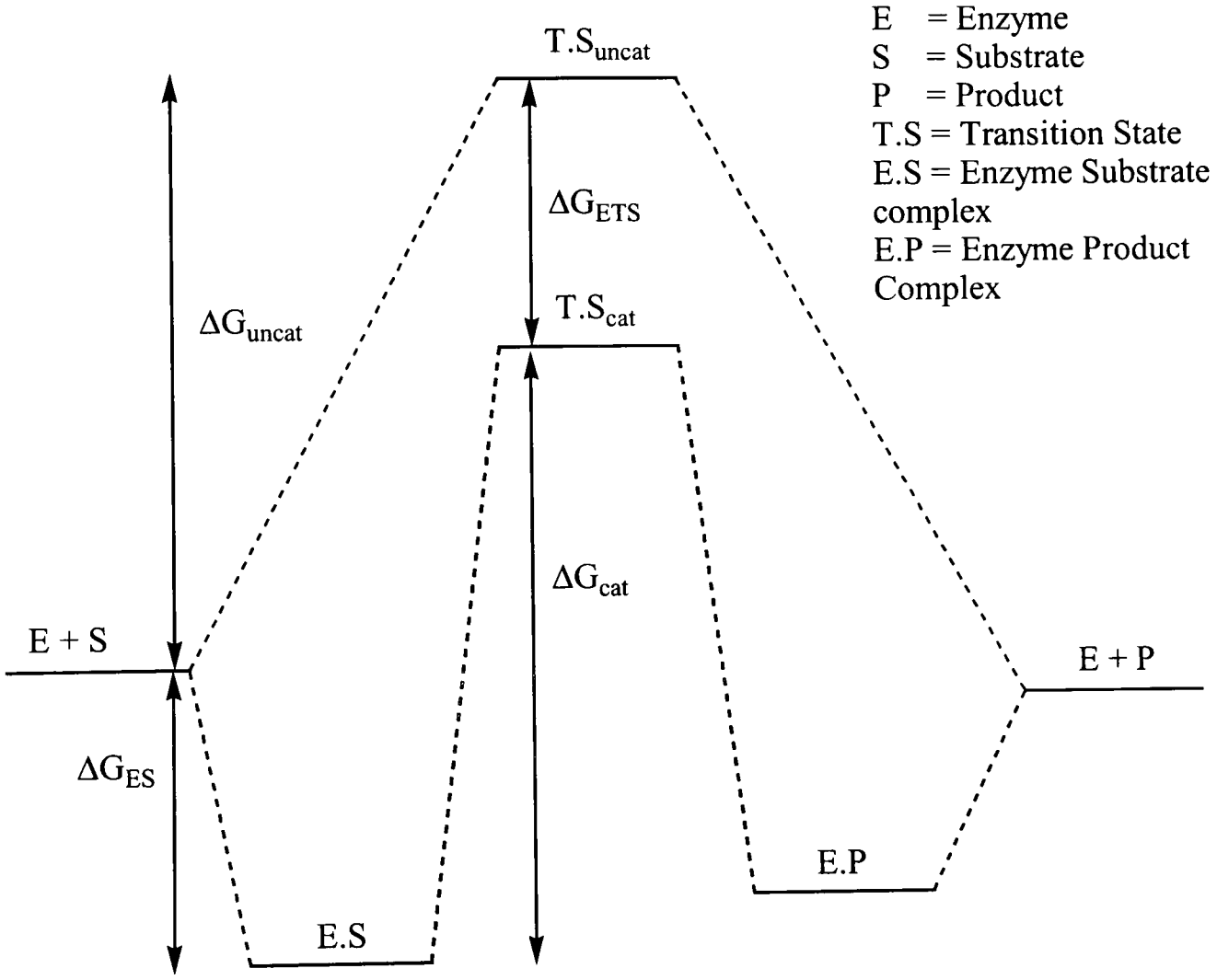


Figure 1

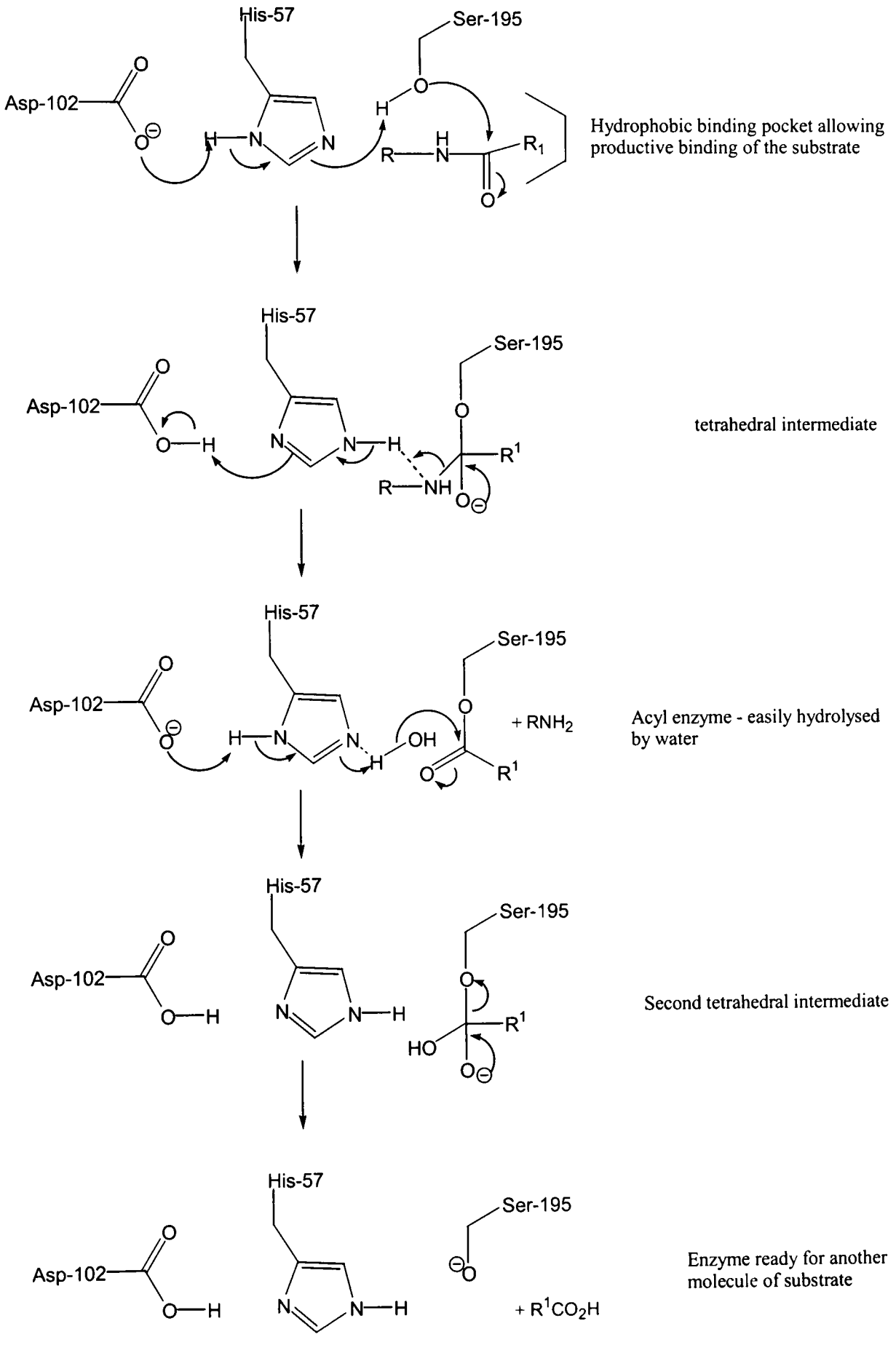
The activation energies for the reaction are denoted by ΔG_{cat} and ΔG_{uncat} for the enzyme catalysed and free reactions respectively. For catalysis to occur, ΔG_{ETS} the free energy of binding of the transition state must be greater than ΔG_{ES}. This results in a lowering of

the activation energy, and thus acceleration of the reaction. For true catalysis to occur however, the system needs to exhibit turnover, i.e. the enzyme substrate complex must be lower in energy than the enzyme product complex. If the opposite is true, then product inhibition of the enzyme is observed. From this picture of enzyme action it can be seen that when designing the active site of an enzyme mimic, we must not only consider transition state stabilisation, but also the thermodynamically favourable release of the product. Clearly, in more complex models, application of the above model of transition state stabilisation is less straightforward.

The design of many early artificial enzymes focused solely on the idea of transition state theory[4]. The above discussion is however a simplification of the real situation, since enzyme catalysis of a transformation often involves an alternative reaction pathway from the non-catalysed reaction, taking advantage of the enzymes ability to reduce the molecularity of multistep sequences. The lack of universal success in the attempt to synthesise an artificial system with the same capabilities as a natural enzyme has meant that we now need to look in more detail at what other factors are involved in enzyme catalysis, and how enzymes achieve their selective binding of the transition state.

Simplistically, there are four main factors involved in enzyme catalysis[5]. First, a chemical apparatus is required at the active site to deform or polarise bonds of the substrate, thus making it more reactive. This can be as simple as hydrogen bonding between the substrate and enzyme, or the binding of a metal. In more complex systems, amino acid residues are present in the active site that are capable of forming covalent bonds with the substrate, usually to form a more reactive intermediate. Secondly, the enzyme must contain a binding site that immobilizes the substrate in the correct geometry relative to the other groups that are participating in the chemical transformation. Thirdly, a correct and precise orientation of the substrate that permits each step of the reaction to proceed with minimal translational or rotational movement about the bonds of the substrate is also required. This is clearly important in controlling the stereochemistry of the reaction. Finally, immobilisation of the substrate must lower the energy of activation of the reaction. All these factors combine to enable the active site of an enzyme to function as a powerful catalyst.

The serine – protease α-Chymotrypsin, provides a beautiful example which incorporates all of the above factors in its mode of action[5,6,7]. The role of this enzyme is to "cut up" proteins on the carboxyl side of aromatic amino acids. The mechanism of hydrolysis is shown in Scheme 1.



Scheme 1

The enzyme contains a hydrophobic pocket, which binds the aromatic residue of the protein, in the correct geometry and orientation for the reaction to proceed. α-Chymotrypsin contains a catalytic triad, consisting of the serine-195, aspartic acid-102, and histidine-57 residues, which interact *via* a charge-relay system. The function of the buried Asp-102 group is to polarise the imidazole ring of His-57, since the buried negative charge induces a positive charge next to it. This allows proton transfer along hydrogen bonds, and the hydroxyl proton of Ser-195 is transferred to His-57. The active serine residue then becomes a reactive nucleophile capable of attacking the scissile peptide bond. An acyl-enzyme intermediate is formed by a covalent bond between the Ser-195 and the substrate, which does not accumulate, but which is readily hydrolysed by water.

This charge-relay system is not the sole factor in the catalytic efficiency of α-chymotrypsin. Stabilisation of the transition state, and thus lowering of the activation energy is achieved by hydrogen bond formation between the carbonyl group on the substrate, and the amide hydrogen of Ser-195 and Gly-193.

This example shows that whilst transition state stabilisation clearly plays an important role in enzyme catalysis, it is not the only factor involved. Therefore when looking to design a new artificial enzyme, we should try to create a mimic that imitates the entire mode of action, and does not just concentrate on one feature.

In order to design an effective enzyme mimic, we also need to have a better understanding of how enzymes achieve their selective binding of the transition state. A number of factors contribute to the difference in free energy of both substrate and enzyme in the unbound (E + S) and bound states (E.S), and between the ground-state (E.S) and transition-state ($E.S^{\#}$) forms of the enzyme substrate complex (Equation 1)[8].

$$E_{solv} + S_{solv} \rightleftharpoons E.S \rightleftharpoons E.S^{\#} \qquad \textbf{Equation 1}$$

The first equilibrium is dependant on specific interactions between enzyme and substrate, comprising Van der Waals forces and hydrogen bonding in the active site, balanced by the enthalpic and entropic contributions from solvation of both

components. In fact, the desolvation of substrate has been considered to be a vital factor. The formal equilibrium between ground state (E.S) and transition state ($E.S^{\#}$) also depends on substrate-enzyme interactions, but is less sensitive to solvation effects since there is little change in the solvent accessible surface at this stage.

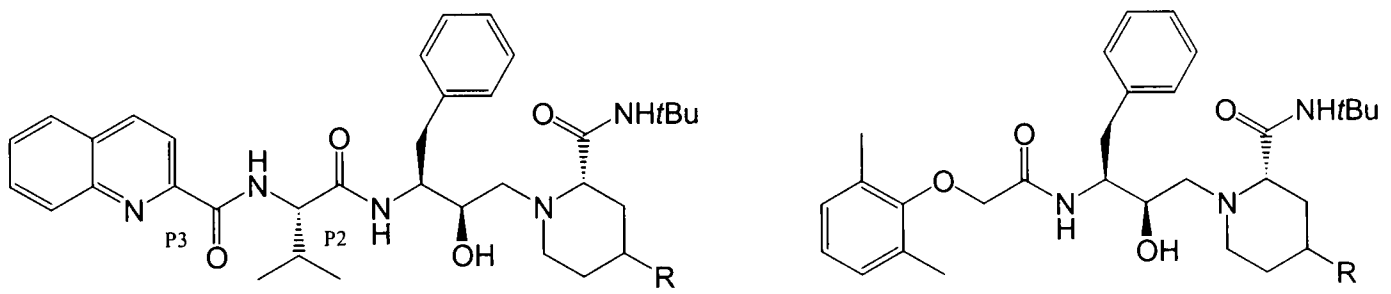
Whilst Van der Waals interactions are significant when calculating the total energy of an enzyme-substrate complex, they do not contribute much to the binding equilibrium, as Van der Waals interactions between elements in the first and second rows of the periodic table are relatively insensitive to the nature of the atoms involved. Therefore there is no significant change on replacing solvent-ligand contacts with solvent-solvent or enzyme-ligand contacts[9].

Electrostatic and hydrogen bonding contribute significantly to the total energy of an enzyme-substrate complex, but as enzymes operate in water, their effects are greatly moderated by solvation. It is often overlooked that formation of a hydrogen bond in a complex is an exchange process and that in order to form this bond similar bonds have to be broken between water and both substrate and enzyme. Likewise new hydrogen bonds are formed between the drug and receptor and between the previously hydrogen bonded water molecules and bulk water. The contribution to the binding equilibrium is the overall energy change involved in breaking and forming all these hydrogen bonds, and generally is found to be a modest value in the order of 1.5kcal mol^{-1} for a neutral-neutral hydrogen bond[10,11,12]. It has therefore been suggested that the role of hydrogen bonding is to determine ligand specificity by creating an energy penalty for binding the wrong ligand.

The contribution of charged hydrogen bonds to binding enthalpy has been found to be more significant. Fersht studied the tyrosyl t-RNA synthase catalysed coupling of tyrosine to ATP to yield tyrosyl-AMP[12]. Site directed mutagensis was used to probe the energetics of the coupling. As the crystal structure of the complex was known, mutations could directly be interpreted as losses in specific hydrogen bonds, salt bridges and hydrophobic contacts. Through a series of point mutations they were able to determine that whilst the neutral-neutral hydrogen bond contributed only 0.5-1.5kcal mol^{-1}, the presence of a charged hydrogen bond contributes up to 4.7kcal mol^{-1}, which is equivalent to a 3000-fold increase in affinity.

More recently the importance of the hydrophobic effect in selective binding has been highlighted. This stabilisation arises from the transfer of a hydrocarbon surface out of water and into a hydrophobic region of an enzyme. The interaction arises not from attraction of the hydrophobic surfaces for each other, but from the favourable change in free energy as ordered water molecules surrounding the hydrophobic surface are released to bulk water. As a result, the effect correlates with the area of the hydrophobic surface that is desolvated. A variety of approaches have been taken to determine the magnitude of this effect, which is worth a minimum of 0.68kcal mol^{-1} or 28calÅ mol^{-1} per methyl group[13, 14].

An example which demonstrates the hydrophobic effect is the binding of two series of HIV-2 protease inhibitors **1** and **2** (Figure 2)[15]. This example also introduces the concept of 'induced fit', in which the receptor undergoes a conformational change in order to optimise hydrophobic interactions with the ligand.



1 **2**

Figure 2

Similar binding affinities are displayed by **1** and **2** despite the replacement of the P2 and P3 substituent in **1** with a phenoxy substituent in **2**. Analysis of the protein-inhibitor complex of **1** by X-ray crystallography revealed hydrogen bonds between the quinoline amide and Asp-29 and Gly-48. The complex with **2** shows induced fit to the dimethylphenoxy group through a shift of 1Å for Asp-29 and a shift of 4Å for the side chain at Asp-30. This elegantly demonstrates two major features of enzyme binding. First, that affinity can be enhanced through hydrophobic interactions, even at the cost of hydrogen bonds. Secondly, it demonstrates the inherent flexibility of enzymes, such that they can adopt different conformations in order to optimise binding.

All of the above effects involve binding between discrete ligand and receptor moieties, and are influential in both of the equilibria shown in equation 1. The changes in active site interactions on going from ground state to transition state have seldom been addressed directly. This is because the system, and hence the bonding interactions, are dynamic in nature. This binding phenomenon, termed "dynamic binding", distinguishes between ground state and transition state recognition. Kirby has used the example of serine proteases to illustrate the concept of dynamic binding[16]. During the transition state for this reaction at least six bonds are being broken or formed (Figure 3).

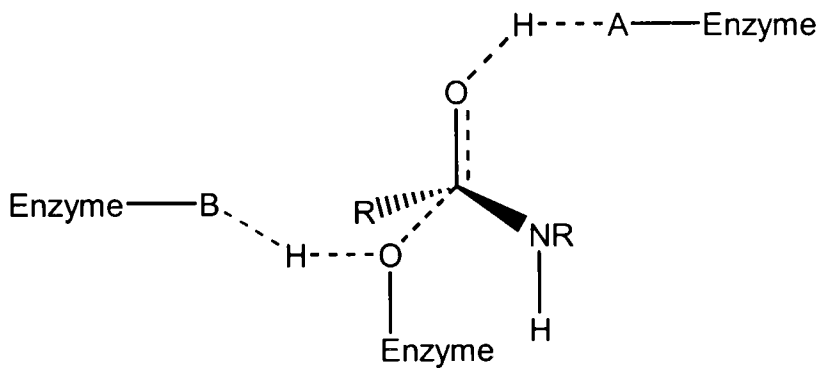


Figure 3

The exact nature of the transition state is unknown, but it is clear that the enzyme uses more than ordinary molecular recognition to "bind" this transition state. "Dynamic binding" interactions are represented by the partially formed covalent bonds at the reaction centre, which have no conveniently modelled ground state counterpart. These interactions play a significant role in the binding and stabilisation of the transition state, and represent one of the major differences between ground state and transition state recognition.

All of the factors described above contribute to the overall binding of the transition state. Combining this information with other knowledge about enzymes mechanisms to create an effective enzyme mimic provides a true test of our understanding of the subject. It is to be hoped that investigations into the design and synthesis of artificial enzymes will not only be aided by this knowledge, but also that these investigations will in themselves lead to a better understanding of molecular recognition and catalysis itself.

3. Previous Approaches to Artificial Enzymes.

Traditional approaches to the synthesis of an artificial enzyme focused on the rational design and synthesis of a complex catalyst molecule. Whilst some impressive results have been achieved in this manner, this type of approach is time consuming, both in conception and practice, and the smallest flaw in design can lead to catastrophic results. More recent strategies have been based on the idea of selection, either by binding, or directly by catalytic activity, and, as we shall see, utilise fields as diverse as molecular biology and polymer chemistry.

3.1 The Design Approach to Artificial Enzymes.

The rational design of macromolecular receptors with appropriately positioned functional groups which try to recreate the binding, catalytic efficiency and microenvironment of enzyme active sites is, of course the most logical and cartisan approach. The functional groups have been chosen to mimic the amino acid residues known to be involved in the enzyme-catalysed reaction. The majority of attention has been focused on the design and synthesis of catalytic cyclodextrins[17-22], but recently some other techniques such as designed cyclophanes[23], and covalent conjugation[24] have shown some impressive results, and proven to be more versatile in the reactions they can catalyse. From the initial concept through to catalytic studies has proven to be a long and arduous process, and although some impressive results have been observed, formation of a real "artificial enzyme", that not only causes a rate acceleration, but also does so in a selective fashion, and with turnover still seems some way off.

3.1.1 Catalytic Cyclodextrins as Enzyme Mimics.

Cyclodextrins are stable, water soluble, cyclic oligosaccharides consisting of 1,4 linked D-(+)-glucopyranose units. This cyclic array produces a hydrophobic cavity in the form of a truncated cone. Primary alcohols are located around the narrow rim, and secondary alcohols around the wide rim (Figure 4).

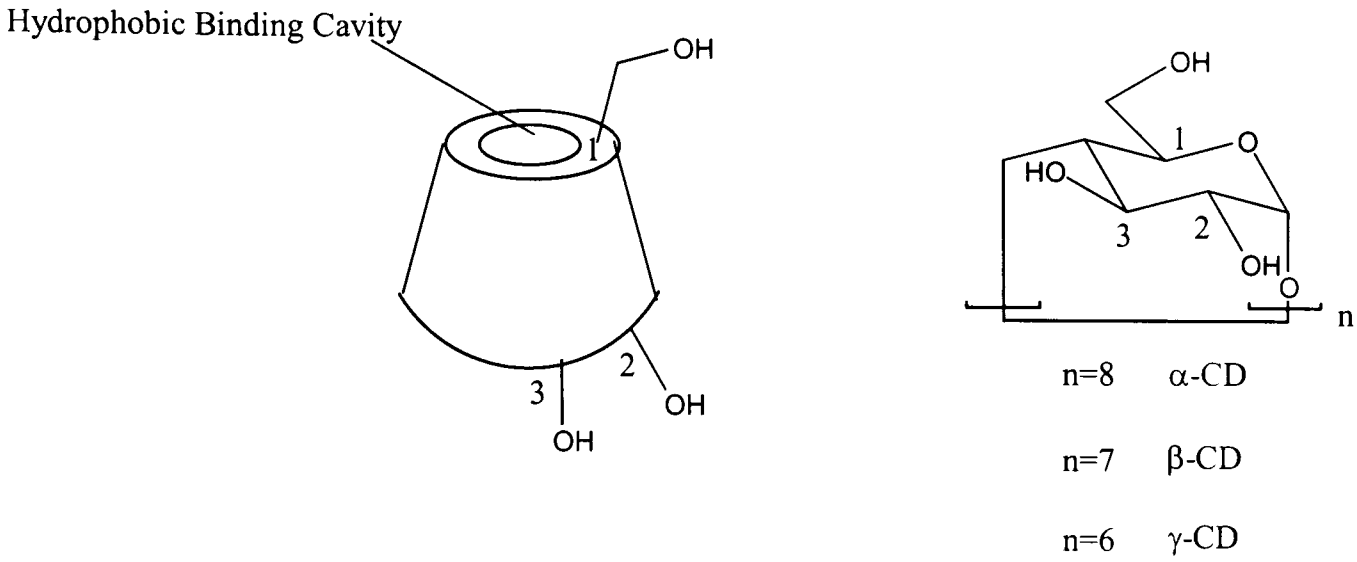
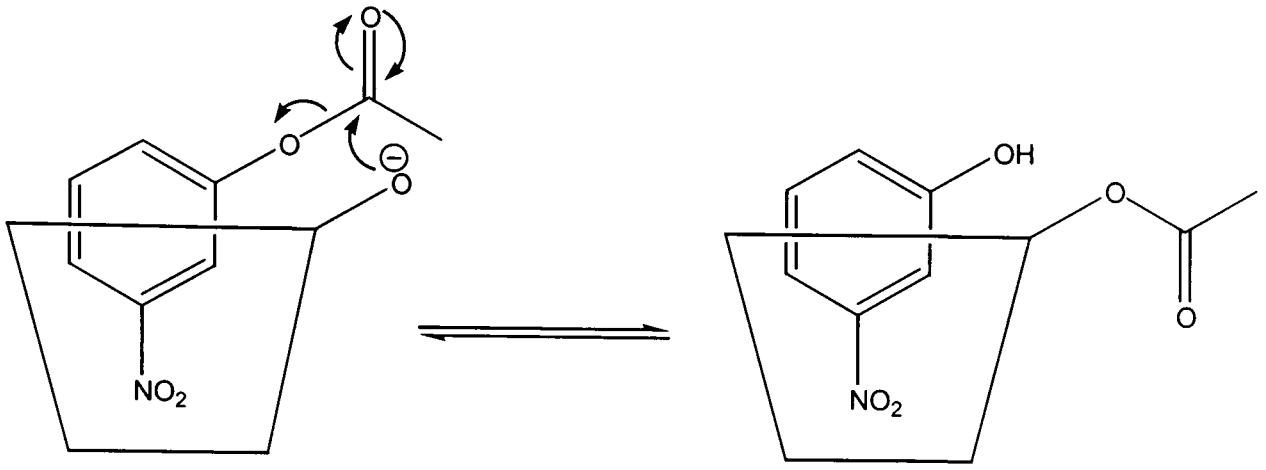


Figure 4

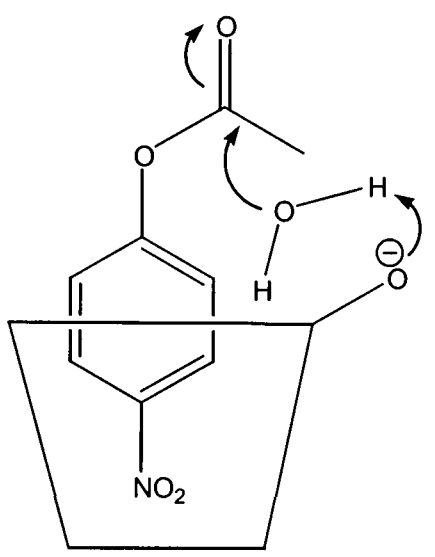
Recognition that the cavity could bind hydrophobic organic molecules in water, and that functionalisation of the hydroxyl groups on the rims was possible, led to cyclodextrins being investigated as potential artificial enzymes.

Unmodified cyclodextrins themselves have been shown to behave as good enzyme mimics, following Michaelis-Menton kinetics in reactions such as hydrolysis of nitrophenolate esters[17,18]. Typically the phenyl group is bound within the hydrophobic binding cavity in such an orientation that nucleophilic attack on the ester by the secondary hydroxyl groups is possible. In the case of *p*-substituted phenyl derivatives however, when the ester group is bound near the central axis of the cavity, reacting groups cannot come into close proximity, and, in this case, general base catalysis competes (Scheme 2).

Significant rate enhancements in the range between 4 and 2150-fold have been observed with unmodified cyclodextrins, but several problems are associated with this type of catalysis[17,18]. First, the hydrophobic phenol product is usually bound tightly within the cavity of the cyclodextrin, and product inhibition is often observed. An excess of cyclodextrin is often used to avoid this problem, but with the unfortunate result that catalytic turnover is not achieved. Second, the nucleophilic mechanism goes *via* an acylated cyclodextrin as an intermediate. The rate determining step of the reaction is now hydrolysis of this intermediate, which in many cases has proven to be less reactive than the original substrate.



Nucleophilic Attack by the Secondary OH



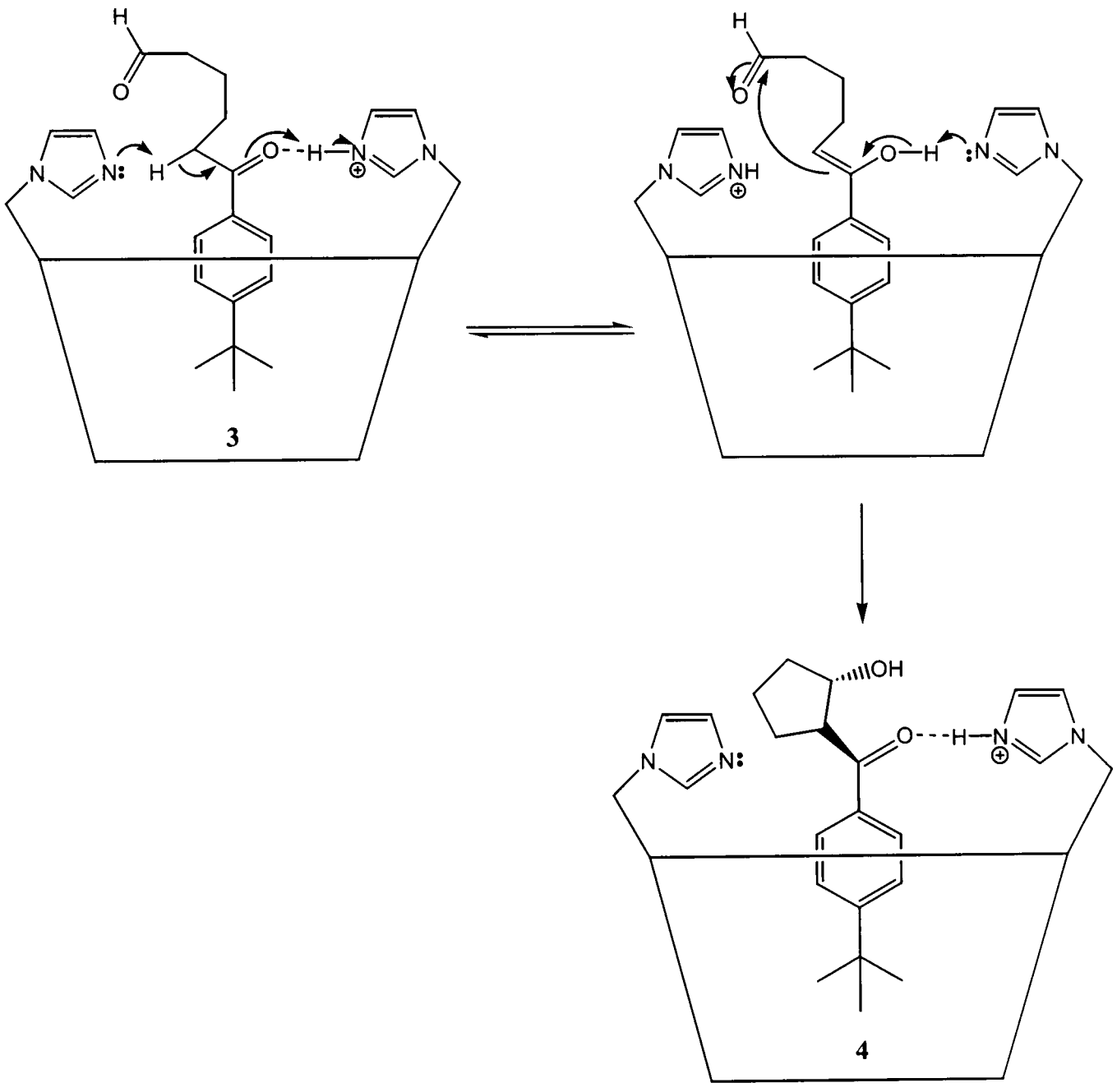
General Base Catalysis

Scheme 2

The close proximity of the hydroxyl groups in cyclodextrins means that functional groups attached at these points are often close enough that they can act in a co-operative manner, such as that observed in the active site of many enzymes. The electrostatic environment in the active site maintains the delicate balance of pKa's required for the various groups to participate catalytically. Histidine is often able to function as both an acid and base in simultaneous bi- or multifunctional catalysis. An example of this is in the enzyme Ribonuclease A, which utilises two histidine residues in the hydrolysis of RNA[25].

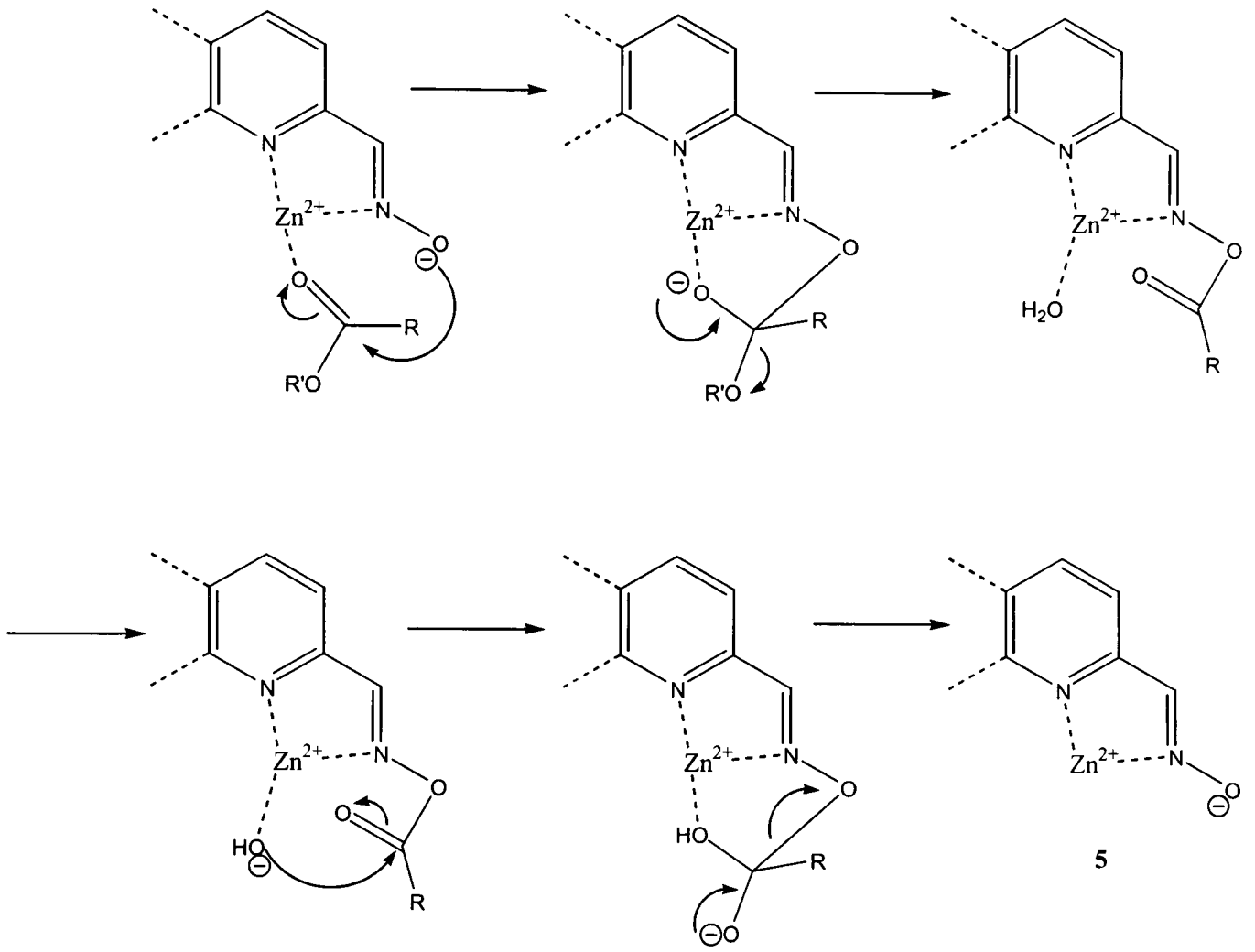
Breslow was inspired by this system to design an artificial aldolase[19]. Two different catalysts were synthesised by arranging two imidazole rings on the secondary rim of β-

cyclodextrin (CD): on glucose units adjacent to each other (AB isomer) and glucose units 154° apart (AD isomer). The intramolecular aldol cyclisation of ketoaldehyde **3** to anti ketoalcohol **4** was studied using these catalysts. In the presence of unmodified β-CD at neutral pH, no significant cyclisation was observed. When the AB isomer was added, a 20-fold rate acceleration was observed relative to the uncatalysed reaction. However, this result was found to be the same as that observed with a monoimidazole functionalised β-CD. With the AD catalyst however, a 50-fold rate acceleration occurred. This clearly indicated that the positioning of the imidazole rings was crucial, and when attached to sugars 154° apart bifunctional catalysis could occur involving both the protonated and neutral imidazole species (Scheme 3).



Scheme 3

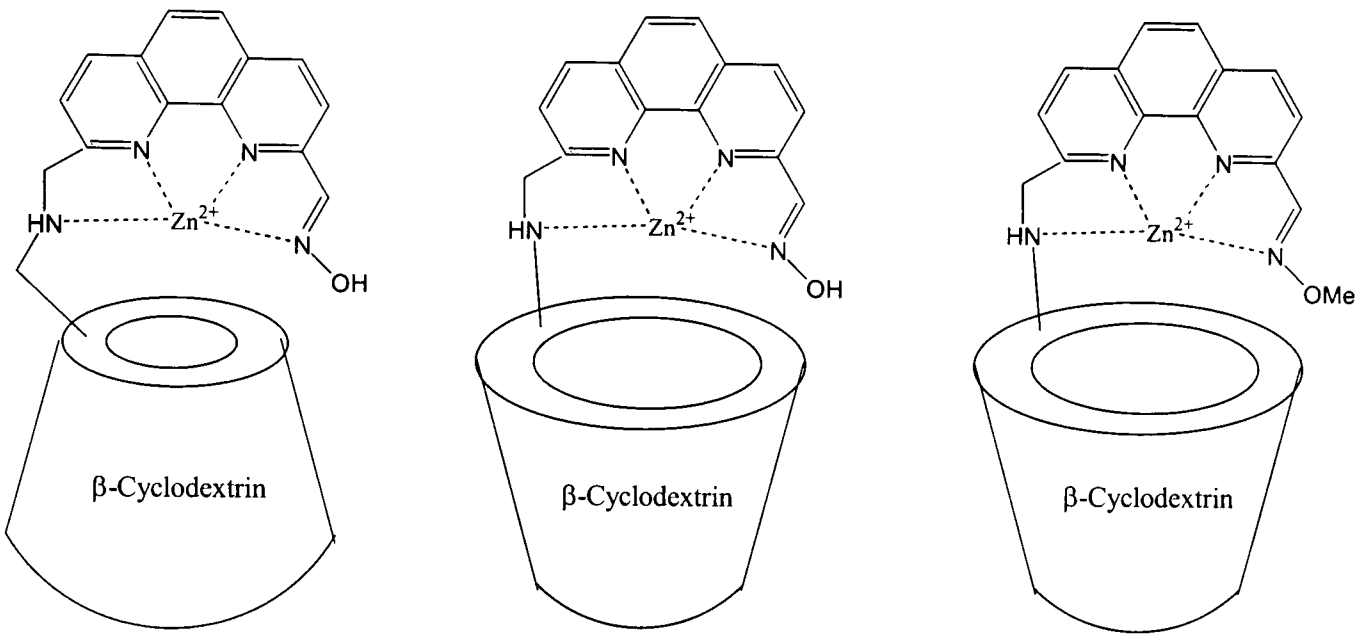
In a more recent example, Breslow noted that the combination of a bound Zn(II) and a coordinated pyridyl-2-carboxaldoxime **5** can be effective in cleaving esters[20]. When the two nitrogens coordinate to the metal, the oxime hydroxyl is acidic, but the oxygen atom is not coordinated and can therefore attack a Zn(II) coordinated carbonyl group. An O-acyl oxime intermediate is formed which can be readily hydrolysed (Scheme **4**).



Scheme 4

Breslow constructed two catalysts based on cyclodextrin binding groups in which a similar oxime was an intrinsic part of the molecules[20]. In one catalyst **6**, the oxime was attached to the primary face, and in the other **7**, the secondary face (Figure 5). When tested as catalysts for the hydrolysis of *p*-nitrophenol acetate in the presence of Zn(II), **7** showed a 22,600-fold rate acceleration over the uncatalysed system. In order to show that the catalysis was a result of the proposed mechanism (Scheme 4), two further cyclodextrin derivatives were synthesised, one with Zn(II) replaced by Cu(II), and the other containing an oxime methyl ether **8**. The rate observed with the Cu(II) catalyst was 250 times slower than the Zn(II) catalyst, and when **8**, containing the oxime methyl ether was used, the rate was only twice that of the uncatalysed reaction. The other catalyst **6** also showed a rate improvement, but only 1500-fold. Breslow thought there

might be two reasons for the greater catalytic activity of the secondary face catalyst. First, there may be a preferential binding geometry, and second the primary face catalyst shows an undesired greater flexibility.

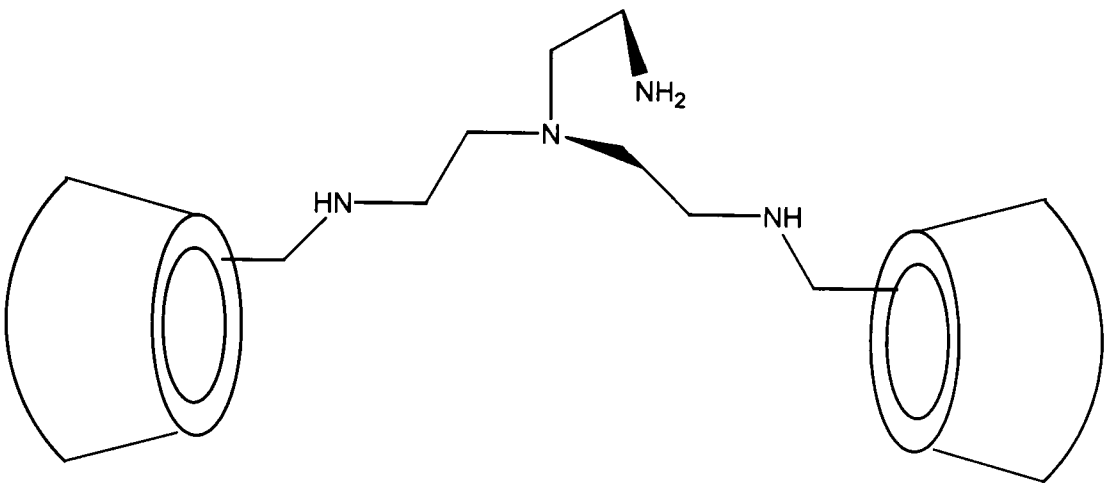


6 **7** **8**

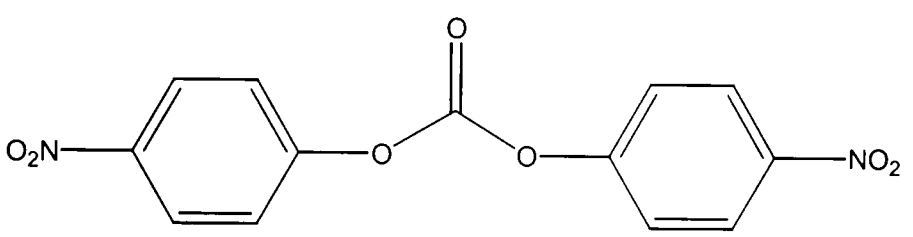
Figure 5

The use of cyclodextrins as enzyme mimics has been extended to incorporate dimers and trimers of cyclodextrins[21,26]. This is a method of introducing more than one binding moiety into the catalyst. In addition, the linker can also contain catalytically active functional groups. The β-cyclodextrin dimer **9** is linked *via* the primary face by tris(2-aminoethyl)amine (tren)[21]. This linker was used as it was thought that at neutral pH, tren would exist primarily as a di-cation. Protonated and neutral amino groups would therefore co-exist, and would open the door to an effective general acid/base catalysis, such as that utilised by a number of enzymes, including ribonuclease A.

The dimer **9** was used as a catalyst in the hydrolysis of bis-*p*-nitrophenylcarbonate **10**, the idea being that the two phenyl groups should sit inside the cavities of **9**, and the carbonate moiety should be in close proximity to the tren amino groups, which could act as acid base catalysts (Figure 6). A 250-fold rate acceleration was observed when a 2.5mM concentration of **9** was added to the reaction. If the pH was increased from 7.5 to 8.5, the rate enhancement was significantly less, thus providing a clear indication that general acid/base, and not nucleophilic action was the catalytic mode of action.



9



10

Figure 6

Traditionally, the use of modified β-cyclodextrins as artificial enzymes has been based on the design approach, and the incorporation of an aromatic ring in the substrate is of course mandatory. However, the subtleties of molecular recognition are difficult to predict *a priori*, and the chemistry for preparing regiospecific β-CD isomers is only in its infancy. Several groups have therefore started to look at combinatorial chemistry, and to combine traditional design, with the more modern advances in selection approaches in their search for the perfect artificial enzyme.

Yu *et al*, have recently described their synthesis of thirteen β-CD libraries, totalling approximately 28,000 compounds in their search for a host with phosphatase like activity[22]. The libraries were generated from reacting β-CD with different combinations of eleven individual amines (**11-21**, Figure 7). The thirteen libraries were then evaluated against controls for their ability to hydrolyse *p*-nitrophenylphosphate (*p*-NPP) in the presence of Zn(II) at neutral pH. The best library, containing amines **11**, **13** and **20**, showed a 66.7-fold increase in rate relative to the reaction carried out only in the presence of Zn(II) and buffer. Although this example does not show the rate

acceleration seen with the more complex designed systems, it is clear that the technique of combining combinatorial chemistry with the rational design of catalytic cyclodextrins holds considerable promise for the future.

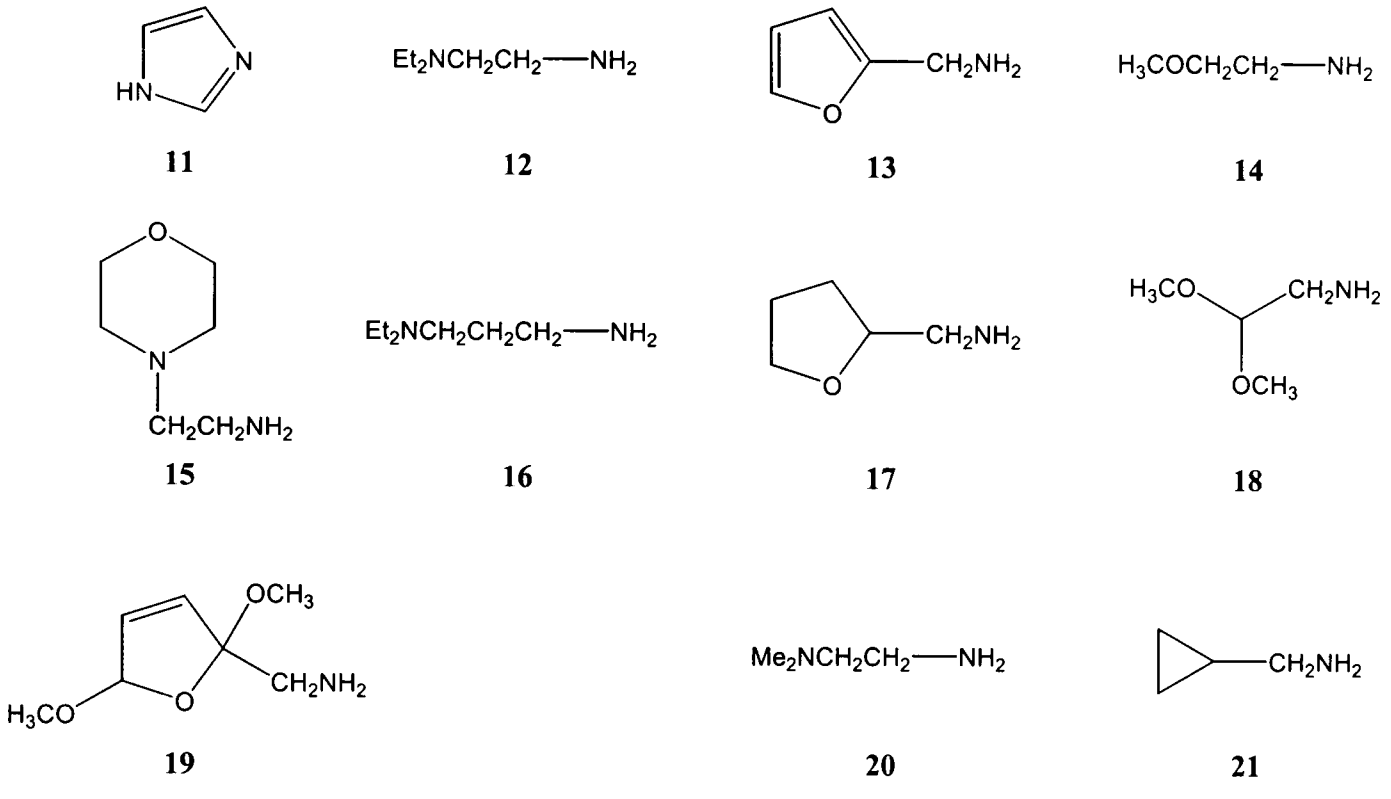


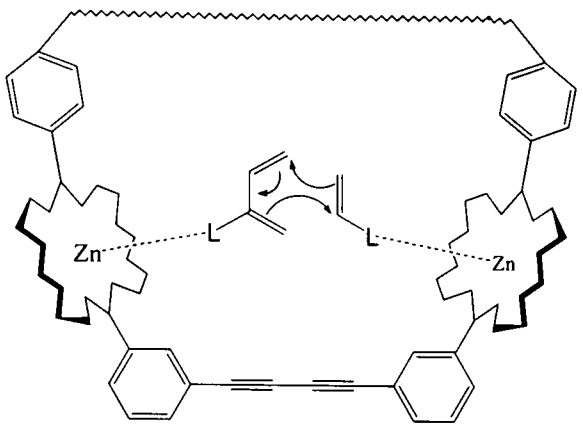
Figure 7

3.1.2 Cyclic Metalloporphyrin Trimers as Artificial Diels-Alderases.

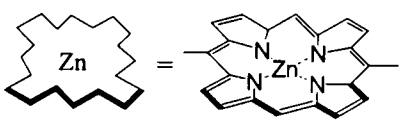
Whilst most effort in the rational design of artificial enzymes has been focused on looking at catalytic cyclodextrins, other synthetic hosts have also been investigated. One problem associated with cyclodextrins is that the size and geometry of the hydrophobic cavity in the host is fixed, and whilst functional groups on the rims can be readily exchanged, reactions that can be catalysed are limited to substrates that can be accommodated within the central cavity. It is therefore desirable to find a method where the cavity can be adjusted to fit to a certain reaction.

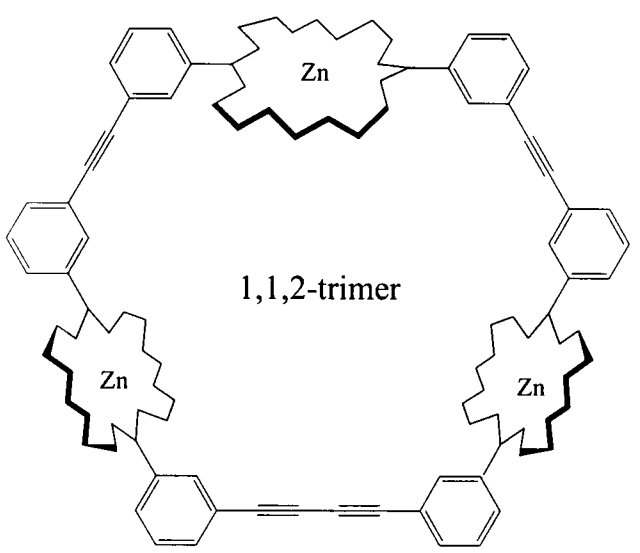
With this in mind, Sanders *et al.* have studied the use of cyclic metalloporphyrin trimers as artificial enzymes[27,28,29]. They have been investigating the catalysis of a Diels-Alder reaction, which is an attractive target for a couple of reasons. First, although a natural Diels-Alderase has now been discovered, there is no catalyst available for synthetic use[30]. Second, there is a considerable challenge involved in controlling bimolecular reactions, since two substrates need to be recognised and correctly oriented at the same time. Their thinking was that porphyrin hosts **22** and **23** (Figure 8) could bind pyridine

base ligands within their cavity, in close proximity to each other, and hence catalyse the reaction simply by virtue of their binding properties (Figure 8)[27,28].

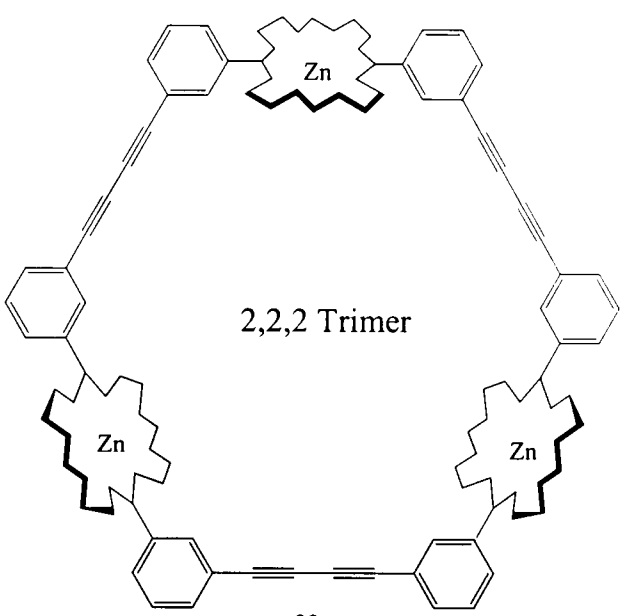


A Porphyrin Artificial Diels Alderase





22

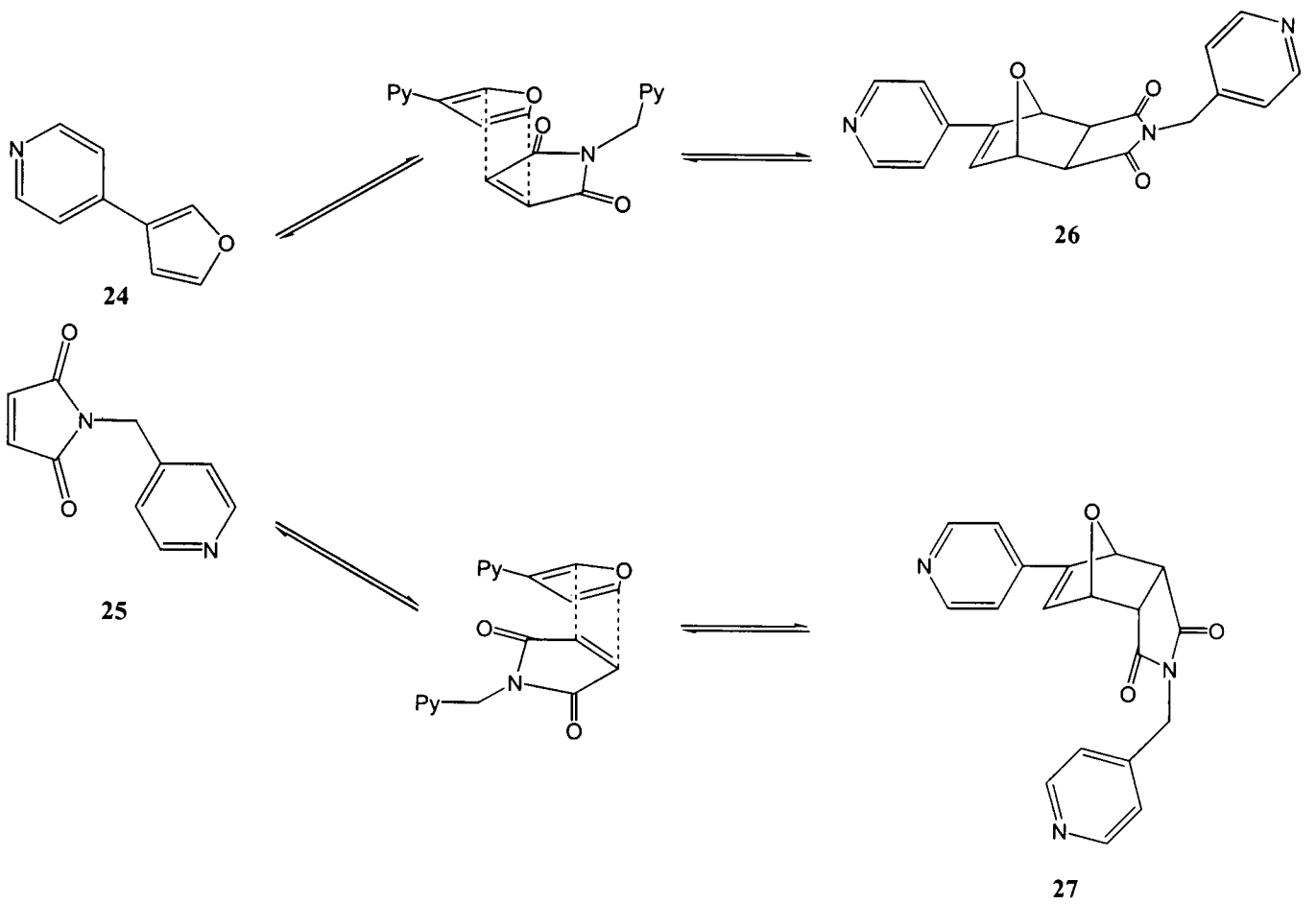


23

Figure 8

The hosts were designed to act as catalysts for the Diels-Alder reaction between the furan-derived diene **24** and the maleimide-derived dienophile **25** (Scheme 5). When the reaction was studied in the absence of hosts, the *exo* and *endo* adducts (**26** and **27**) were produced in an approximately 2:1 ratio, with the *exo* adduct **26** being the thermodynamic product. At 30°C, in the presence of one equivalent of the 2,2,2 trimer **23**, the *exo* adduct **26** was the exclusive product, and a 1000-fold rate acceleration was observed. When the 1,1,2 trimer **22** was used however, the selectivity was reversed, with only the *endo* adduct **27** observed, with a 500-fold rate acceleration. The reversal of stereoselectivity between the two cyclic trimers was thought to be the result of the greater flexibility of the 2,2,2 trimer **23**. Neither of the hosts contains the ideal geometry to bind the *exo* adduct's transition state, but the larger trimer is more flexible, and therefore able to adjust to meet the geometrical constraints of the *exo* pathway. This stereochemical reversal in its own right shows significant progress in the rational design

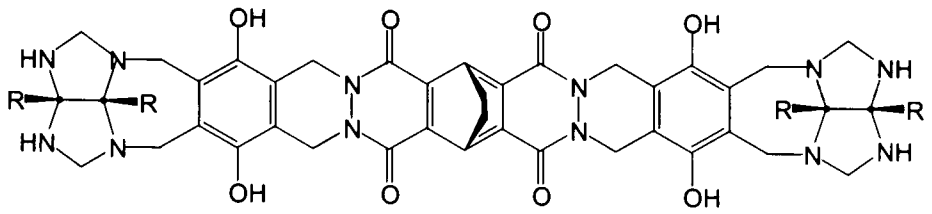
approach. It is worth noting however that the host always had to be used in a stoichiometric amount, since product inhibition prevented catalytic turnover, and this still poses one of the major challenges for the rational design of these systems.



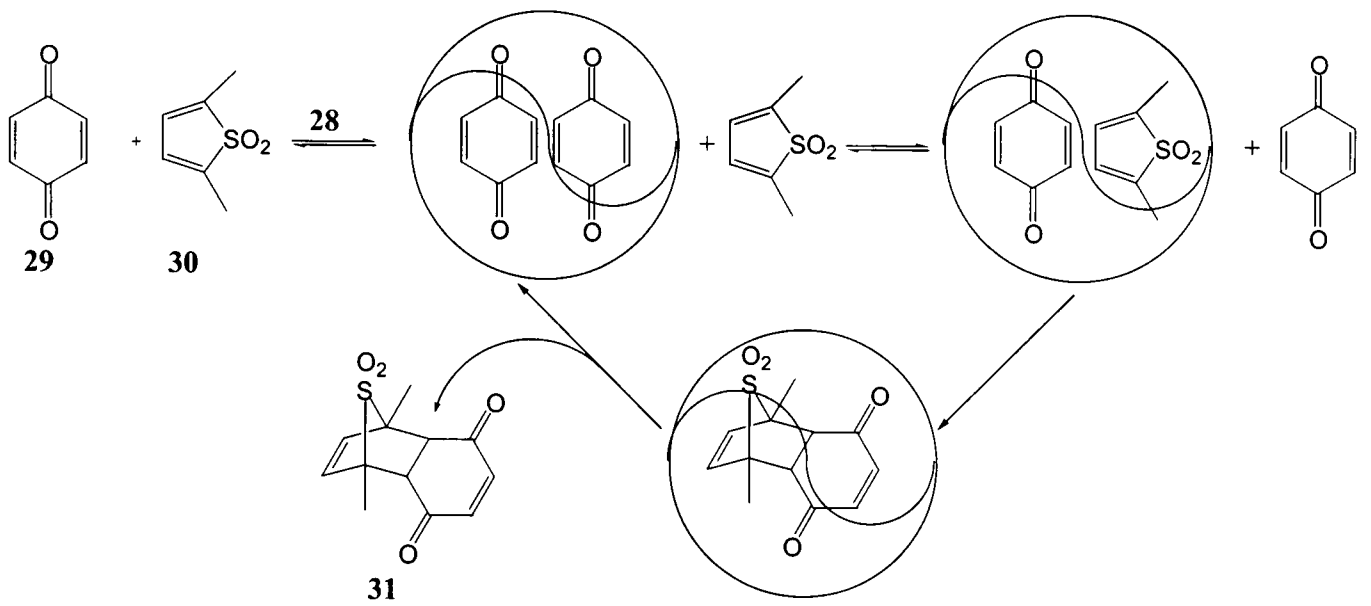
Scheme 5

3.1.3 Self-Assembled Molecular Capsules as Catalysts.

Rebek has also investigated artificial enzymes which are capable of catalysing the Diels-Alder reaction[31]. Previous work had focused on the synthesis of reversibly assembled self-dimers of the multi-ring structure **28**[32]. The dimer is held together by sixteen hydrogen bonds, and has a pseudo-spherical shape, which has led to it being labelled as a "hydroxy-softball". As hydrogen bonds are dynamic, the "softball" forms and dissipates on the timescale of milliseconds, and provides a temporary receptacle for smaller complementary molecules. Observation that the cavity formed can accommodate two molecules of solvent benzene raised the possibility that the capsules could be used as chambers for bimolecular reactions. Studies focused on the specific Diels-Alder reaction between *p*-benzoquinone **29** with the thiophene dioxide derivative **30** in *p*-xylene (scheme 6)[31].



28 R=4-*n*-heptylphenyl



29 **30** **31**

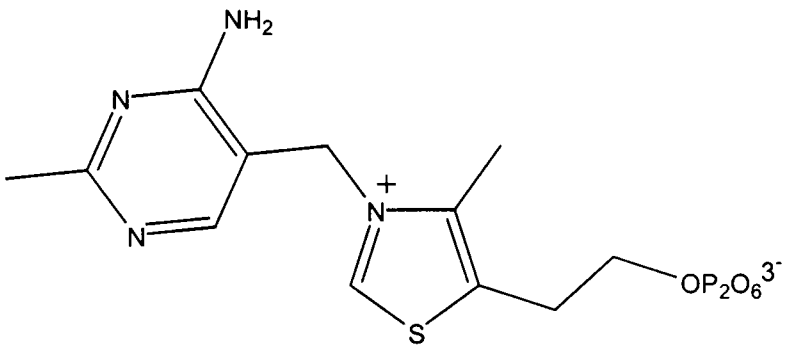
Scheme 6

Previous efforts using a different substrate had found that whilst the reaction showed good rate acceleration in the presence of dimer **28**, product inhibition proved to be problematic[33,34]. In this case however, it was hoped that extrusion of SO_2 from the adduct **31** would leave a product whose dimensions could no longer be accommodated by the dimer. In fact, the adduct **31** itself proved to be an unwelcome guest in the cavity of the dimer, and was driven out by the *p*-benzoquinone **29**, thus allowing the system to show catalytic turnover. Rate increases with this system were more disappointing however, showing only a modest 10-fold acceleration when compared to the background system.

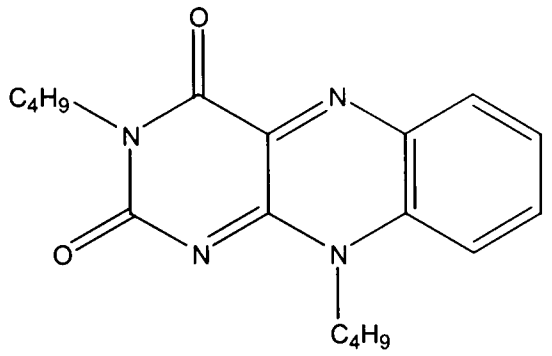
3.1.4 Catalytic Cyclophanes.

Diederich has pursued the mimicry of the mode of action of pyruvate oxidase, which utilises thiamine diphosphate **32** (ThDP) and flavin adenine dinucleotide **33** (FAD) as cofactors to catalyse the reaction from pyruvate to acetyl phosphate[23]. In a similar system, aldehydes can be oxidised to carboxylic esters in an alcohol solvent by a thiazolium ion in the presence of flavin (Scheme 7). ThDP-mediated decarboxylation of

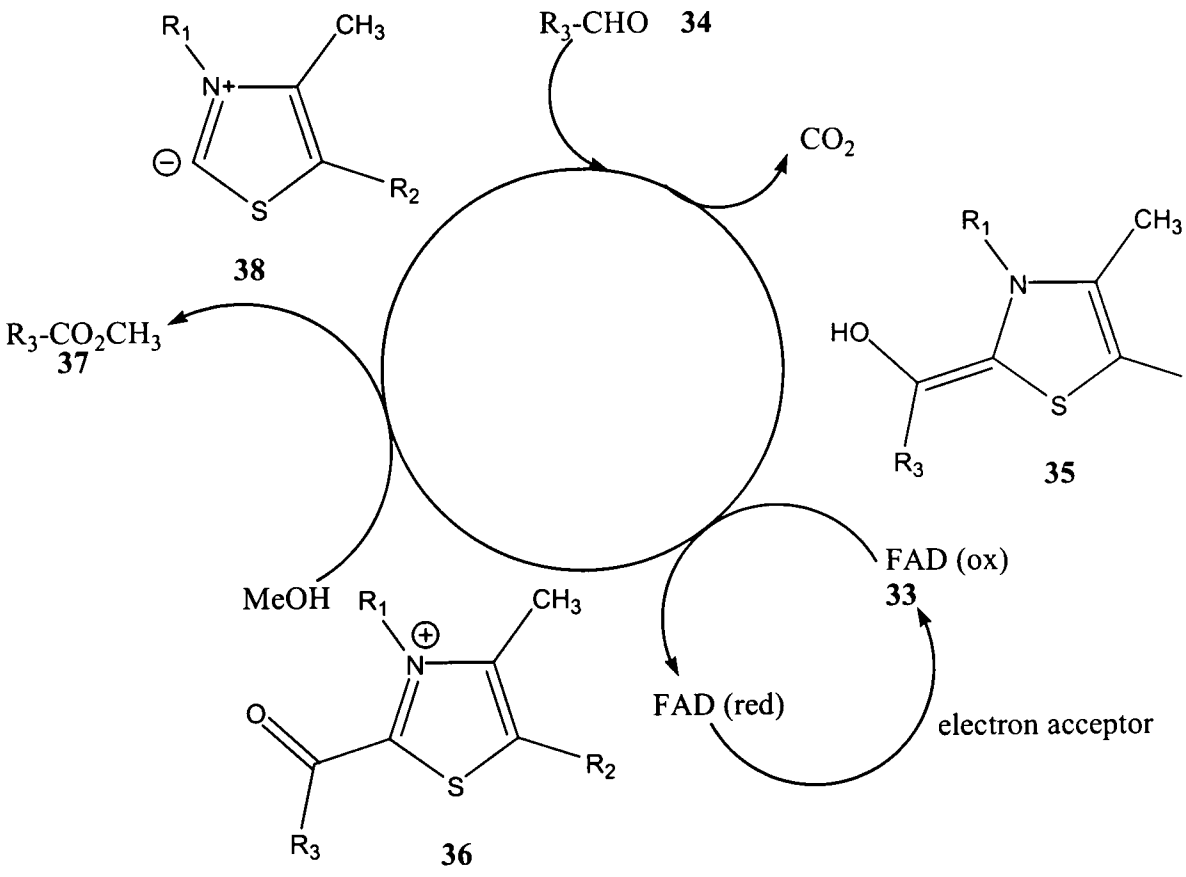
the aldehyde **34** generates an active aldehyde **35**, which is oxidized by FAD **33** to form a 2-acetylthiazolium intermediate **36**. This reacts with methanol to produce the ester product **37** and regenerate the thizolium ylide **38**.



ThDP **32**



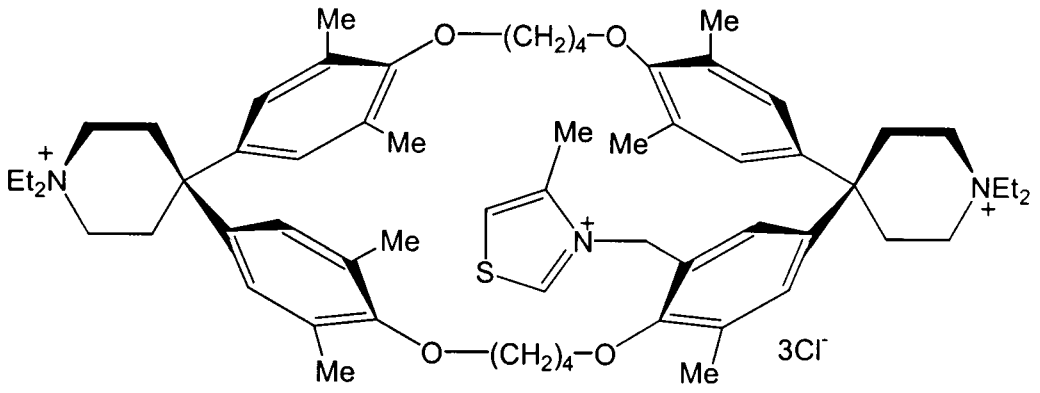
FAD **33**



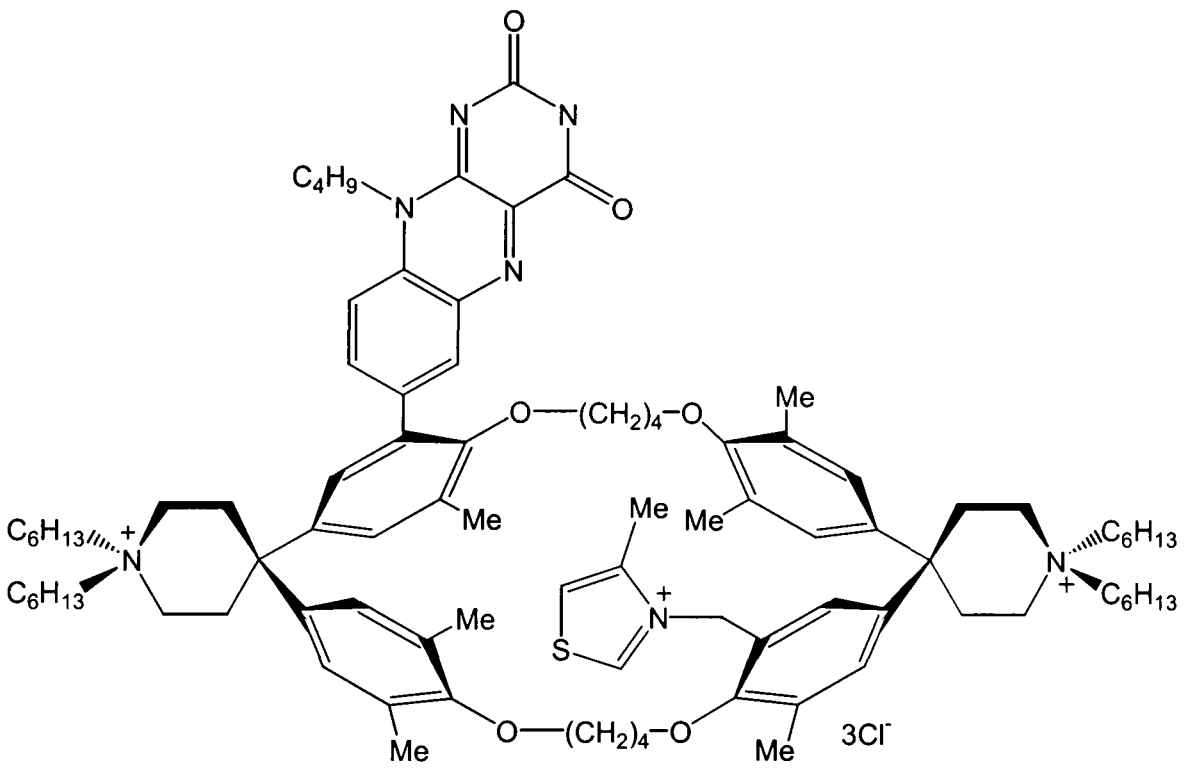
Scheme 7

Initial work used the synthetic thiazolio-cyclophane **39** (Figure 9) as a catalyst, which was designed to contain a binding site for aromatic substrates[35]. In methanolic solution, **39** catalysed the oxidation of an aromatic aldehyde, but a flavin derivative had to be added in order for the reaction to proceed. A molecular modelling study showed that intermolecular oxidation by the flavin was hindered, as the active intermediate bound

within the cavity was not readily accessible. Diederich therefore proposed that cyclophane **40** should be a better catalyst as the flavin moiety was now covalently incorporated into the catalyst[23].



39

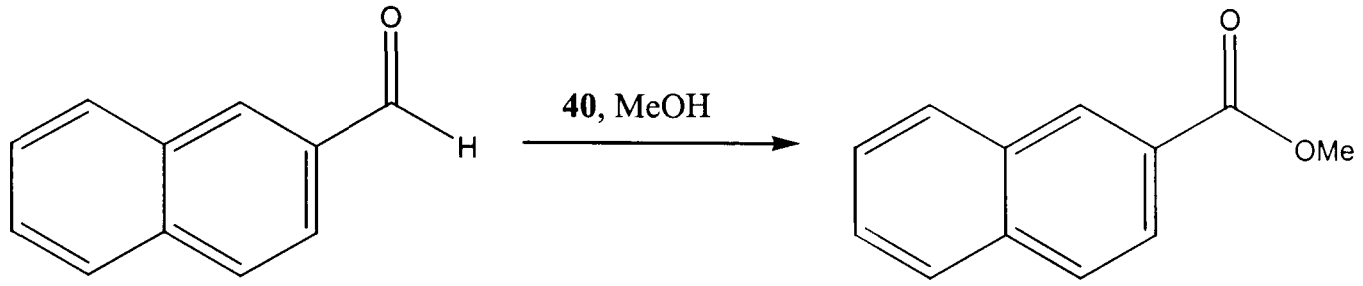


40

Figure 9

The modified cyclophane **40** was indeed found to catalyse the transformation of naphthalene-2-carbaldehyde **41** to methyl naphthalene-2-carboxylate **42** (Scheme 8) in basic methanol, with a k_{cat} $0.22s^{-1}$. This represented a 20-fold rate increase over the transformation using catalyst **39**. In order for the system to show true catalysis, the flavin moiety needs to be reoxidised *in situ*. This was achieved by regeneration at a working electrode potential of –0.3V vs Ag/AgCl[23,36]. Under these conditions, this

system represents one of the most efficient enzyme mimics yet produced, with a turnover of up to 100 catalytic cycles.



41 **42**

Scheme 8

Diederich then tried to go one step further in his quest for the perfect artificial enzyme[37]. The transition states in the formation of the intermediate active aldehyde in the reaction is of a reduced polarity compared to the ground state substrate. The active site of the enzyme is in a microenvironment of reduced polarity, and this is beneficial in catalysing the reaction. He hoped to reproduce this microenvironment by inserting thiazolium-cyclophanes into the core of globular dendrimers, to form catalytic dendrophanes **43** and **44**.

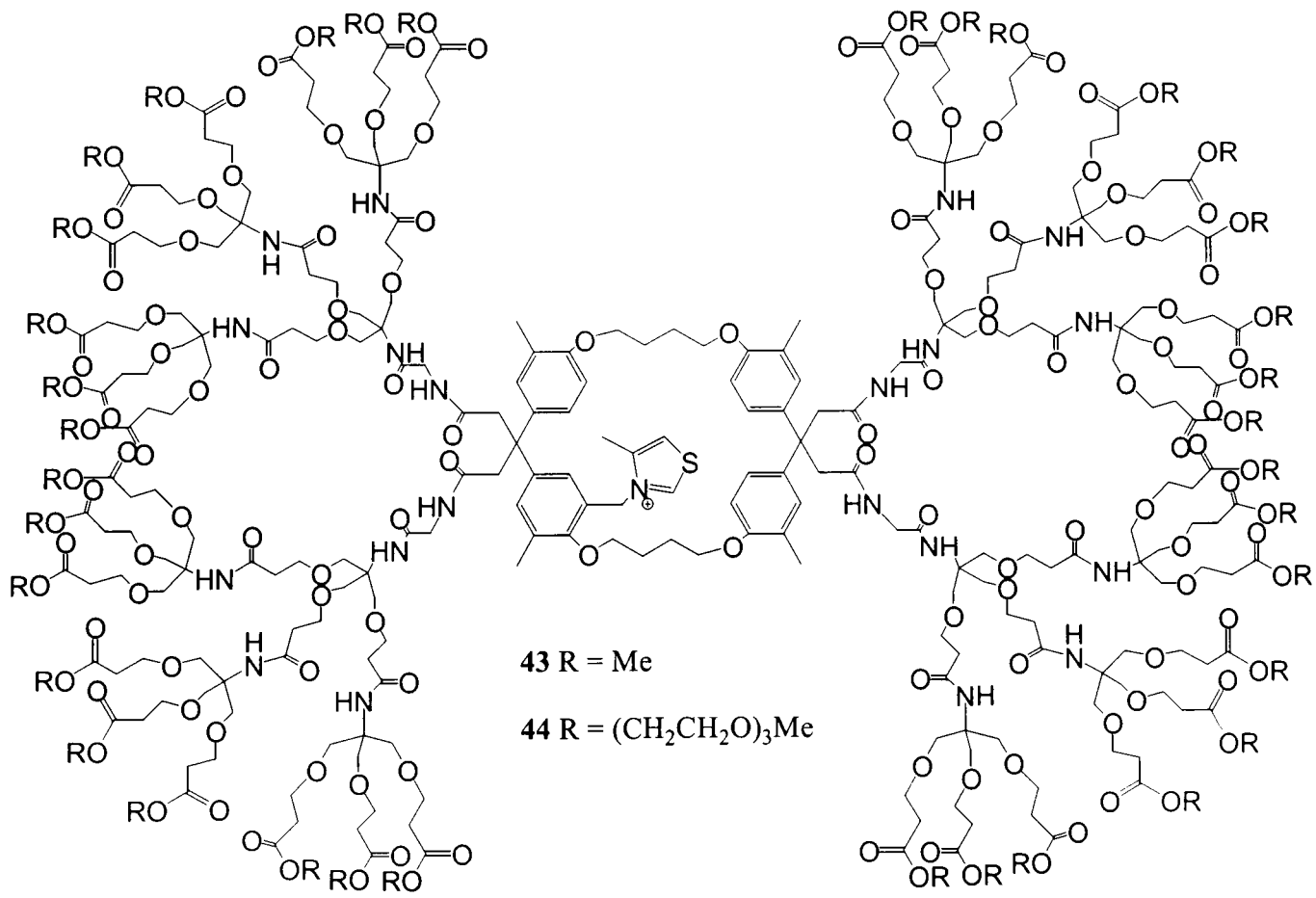
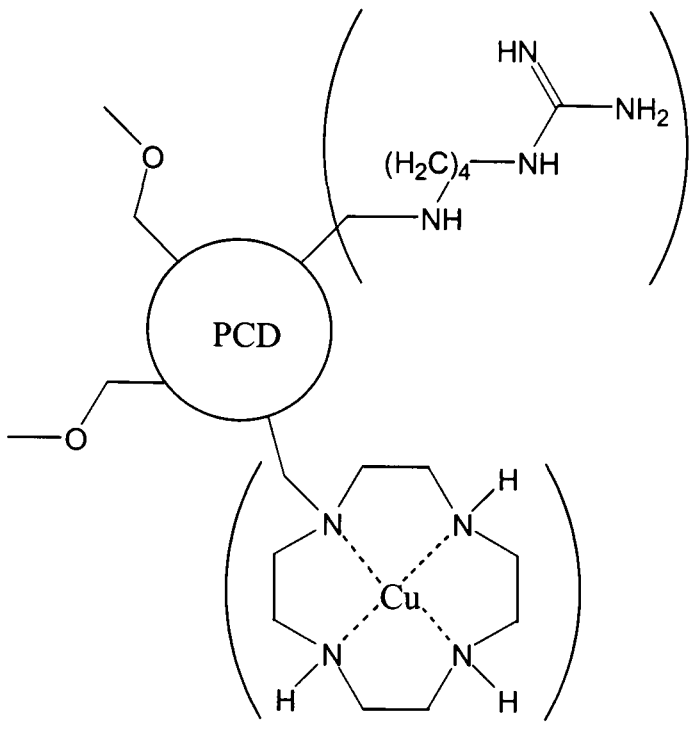


Figure 10

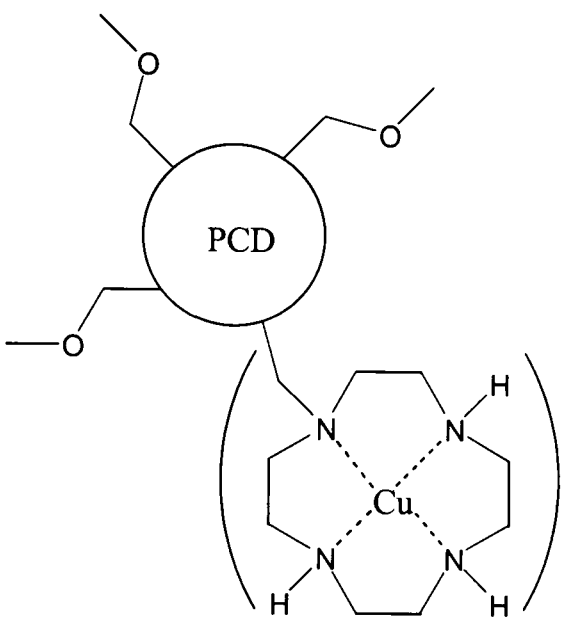
When dendrophanes **43** and **44** were tested for catalysis on the naphthalene-2-carbaldehyde system, an unexpected decrease of between 50 and 160-fold rate acceleration was observed when compared with the reaction catalysed by cyclophane **40**. This posed the question as to why the dendrophanes showed such reduced catalytic activity despite offering an active site with good binding, catalytic groups, and a favourable microenvironment of reduced polarity. Diederich reasoned that the intermolecular electron transfer from the active aldehyde intermediate to the flavin additive may have become rate determining due to the presence of the sterically shielding dendritic branches. These dendritic branches may also affect other transition states in the reaction, and mask any favourable contributions from the reduced polarity found within the active site. The concept behind this proposal was ingenious, but again shows the limitations of synthesising artificial enzymes by rational design. The protein shell around an enzyme site not only ensures the correct positioning of the catalytically active residues and provides the appropriate polarity, but also plays a vital role in keeping the site free of steric hindrance.

3.1.5 Polymeric Enzyme Mimics Designed with a Catalytically Active Site.

Polymers have been used as backbones for catalysts since the 1970's[38]. If a specific binding site is attached onto the backbone of the polymer, then an artificial enzyme can be created. The guanidinium group is utilised by many enzymes to help with recognition of specific substrates, e.g., the guanidinium ion of Arg-145 in carboxypeptidase A plays an essential role in recognising the carboxylate anion of the substrate. Suh attempted to mimic the action of this enzyme by randomly attaching guanidinium ions, and cyclen groups to poly(chloromethylstyrene-co-divinylbenzene) (PCD) (Figure 11)[39]. The role of the cyclen was to complex a metal ion. It was thought that the guanidinium moiety should be able to bind the substrate in an orientation such that the metal ion could perform the catalysis. A Cu(II) ion was then complexed into the cyclen to form polymer **45**. As a control, another derivative **46**, of PCD was formed which contained only the cyclen complex.



45

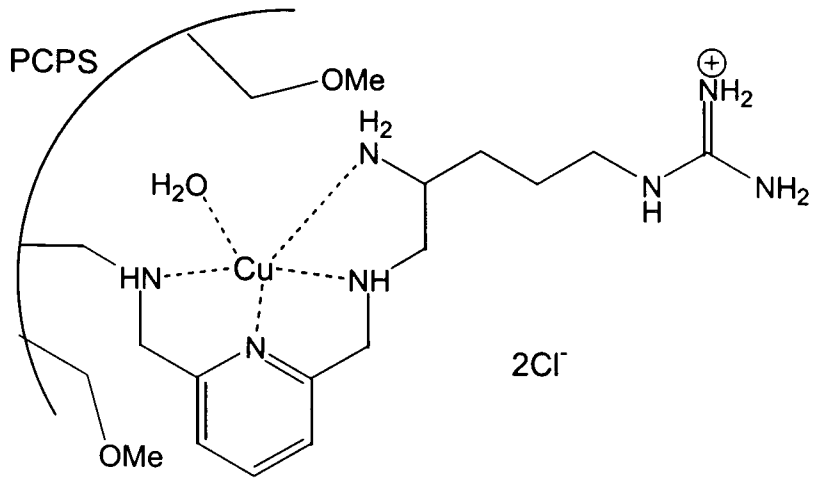


46

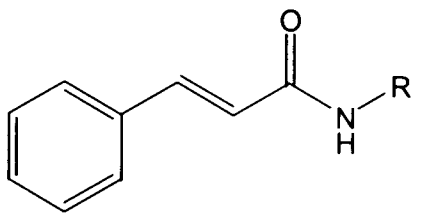
Figure 11

The catalytic activity of the derivatives was tested on the hydrolysis of both the light (MW 25kDa) and heavy (MW 50kDa) chains of bovine serum γ-globulin. Kinetic data showed that in the presence of derivative **45**, containing both the guanidinium group, and Cu(II) cyclen complexes, the half-life of γ-globulin was as short as 40mins. This is 10^8 times shorter than that of the spontaneous hydrolysis of peptide bonds. The activity shown by derivative **45** is 10^4 times greater than that exhibited by the free Cu(II) cyclen complex. This enhancement is likely to be caused by the microenvironment surrounding the Cu(II) complex in the gel-like phase of the polymer, which in similar fashion to enzyme active sites contains both hydrophobic and regions. Derivative **46** also catalysed the reaction, but was 12-15 times less active than **45**.

Whilst the polymer **45** showed good catalytic activity towards the hydrolysis of large proteins, it is relatively inactive towards small amides. Suh reasoned that this was probably due to the random attachment of the functional groups creating a large active site within which the catalytic groups are not in sufficiently close proximity to each other[40]. Accordingly, in order to create a more efficient mimic of metallopeptidases, the guanidinium ion was covalently attached to a Cu(II) tetra-aza complex and then incorporated into a partially chloromethylated cross-linked polystyrene (PCPS), with the result of providing a more well defined active site **47** (Figure 12)[40].



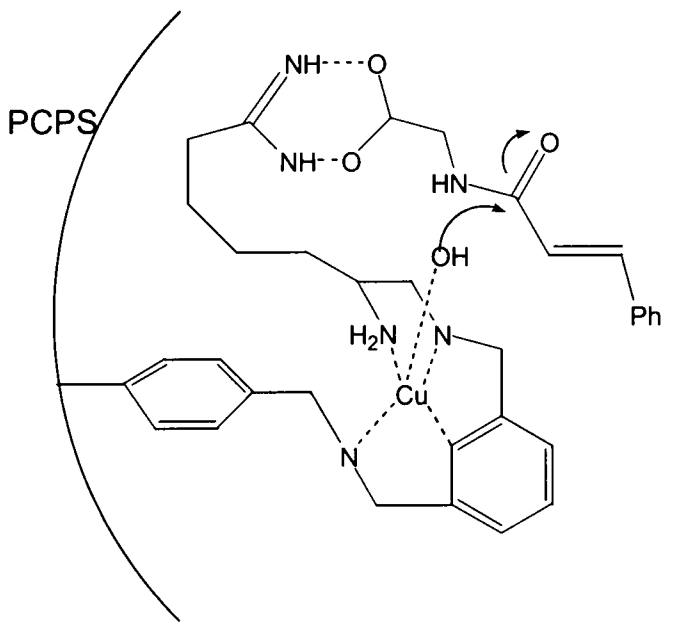
47



48 R = CH_3
49 R = CH_2CO_2H
50 R = $CH_2CH_2CO_2H$
51 R = $CH_2CH_2CH_2CO_2H$

Figure 12

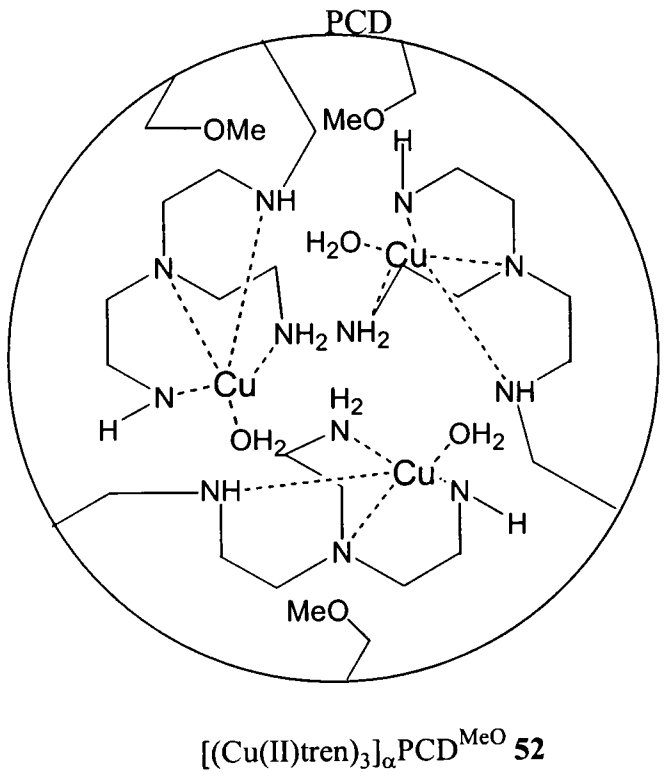
The hydrolysis of amides **48-51** was studied in the presence of polymer **47**. The hydrolysis of amide **48** was not affected by the addition of **47**. On the other hand, hydrolysis of amides **49-51** showed a rate increase in the presence of **47**. For example k_{cat} for **49** was greater than 0.06 h^{-1}, which can be compared with k_o of 2 x $10^{-6}h^{-1}$ for the spontaneous hydrolysis of unactivated small amides. All of the amides hydrolysed contain carboxyl groups, which led Suh to propose the mechanism shown in Scheme 9. Thus, the carboxylate group is bound by the guandinium ion in the correct orientation and geometry for the bound hydroxide ion to make a nucleophilic attack on the carbonyl centre. A similar approach has been used to create cobalt complexed polymers capable of hydrolysing DNA[41].



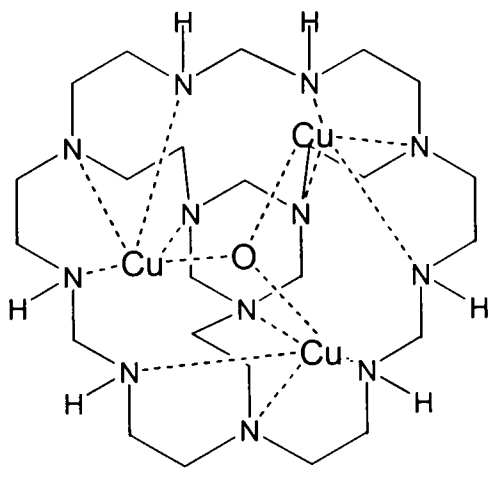
Scheme 9

3.1.6 Artificial Enzymes built by Transfer of Catalytic Elements Combined in a Prebuilt Cage.

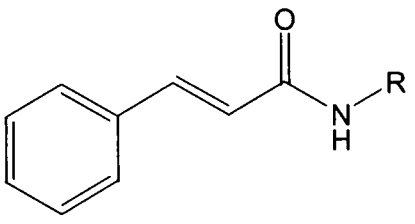
Enzymes such as methionine aminopeptidase, metallo-β-lactamase, and proline dipeptidase catalyse the hydrolysis of acyl derivatives by collaboration between two or more metal ions in the active site. If a receptor site could be designed with two or more metal ions in close proximity, then it seems reasonable to suggest that the receptor should be able to function as an artificial enzyme. Suh prepared a trinuclear artificial metalloenzyme **52** through transfer of three moieties of tris(2-aminoethyl)amine (tren) from the trinuclear macroyclic Cu(II) complex **53** to PCD (Figure 13)[42]. Active sites containing varying quantities of the tren moieties were also prepared.



$[(Cu(II)tren)_3]_\alpha PCD^{MeO}$ **52**



53

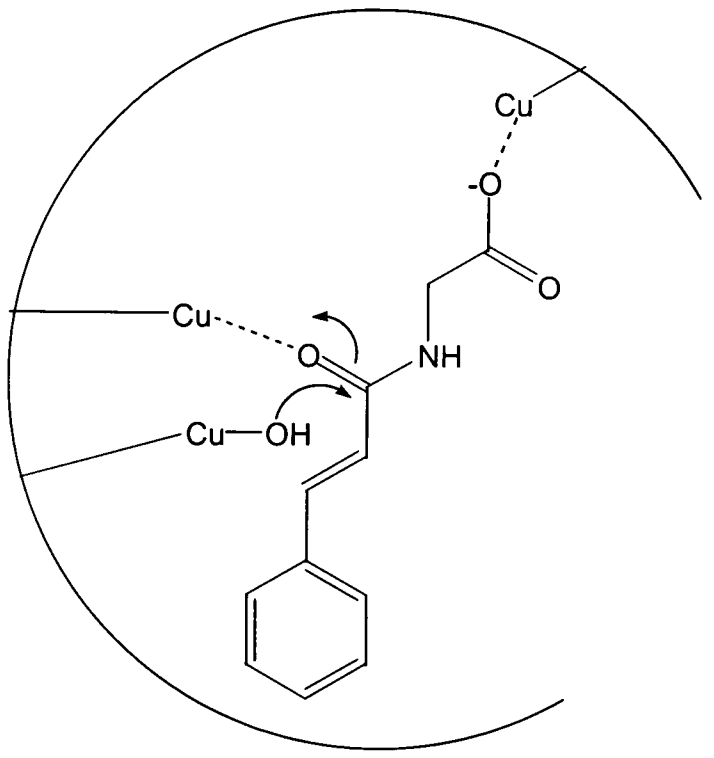


48 R = CH_3
49 R = CH_2CO_2H
50 R = $CH_2CH_2CO_2H$
51 R = $CH_2CH_2CH_2CO_2H$

Figure 13

The catalytic activity of **52** was tested on the hydrolysis of unactivated amides **48-51**. When the neutral amide **48** was used as the substrate, no hydrolysis was observed after 54 hours at pH8 and 50°C. All three of the carboxyl containing amides **49-51** were

however effectively hydrolysed by **52** under the same conditions. The catalyst **52** containing three tren moieties was found to be the most effective catalyst, although all of the catalysts showed much higher catalytic activity than other previously reported synthetic peptidases. The mechanism of action was proposed to be that shown in Scheme 10, with the carboxyl group coordinated to one of the Cu(II) ions, and a second Cu(II) ion coordinating the amide oxygen. The third Cu(II) ion possessing a hydroxyl group is then considered to be in the ideal geometry for nucleophilic attack on the amide.



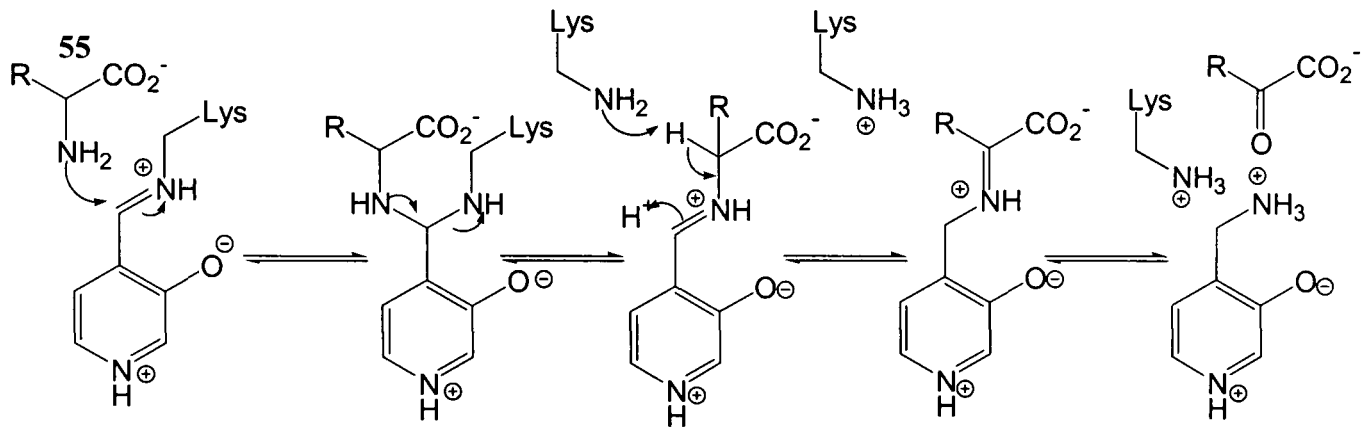
Scheme 10

3.1.7 Chemogenetic Design of Artificial Enzymes.

Semisynthetic enzymes such as thiazol-papain[43], and seleno-subtilisin[44] have been designed and synthesised by the covalent conjugation of a catalytically active group to a protein scaffold. The protein mimics the active site of an enzyme by enhancing binding, and providing a chiral environment. Until recently however, the interaction of functional groups from the protein scaffold with the cofactor or substrate had not been investigated.

Pyridoxamine phosphate is utilised as a cofactor by enzymes that catalyse transamination reactions[45]. The mechanism of action is shown in Scheme 11. The aldehyde form of the cofactor is covalently bound *via* a Schiff base linkage to the ε-

amino group of a lysine residue **54**. Nucleophilic attack on this imine **54** by the amino group of an incoming substrate **55** releases the lysine, and an aldimine complex **56** is formed between the substrate and aldehyde. The same lysine residue then acts as a general base in the subsequent conversion of the aldimine **56** to an internal ketimine **57**. Hydrolysis yields the pyridoxamine **58** and products.



Simplified Transamination Half Reaction

Scheme 11

Häring and Diestefano combined the covalent conjugation approach with site-directed mutagenasis to construct a transaminase active site containing a lysine residue capable of Schiff base formation and general acid/base catalysis[24]. Intestinal fatty acid binding protein from rats (IFABP) was used as the protein scaffold as its spherical shape was thought to be beneficial in correctly positioning functional groups in a 3-dimensional fashion. A cysteine residue was introduced into the protein at residue 60 *via* mutagenasis, and pyridoxamine was tethered to this residue through a sulfide bond. Modelling was used to identify appropriate sites to position the active lysine residue. The ideal positions were found to be residues 38 and 51, and site-directed mutagenesis was used once again to introduce the lysine residues to IFABP at these positions. The two conjugates were then tested for their ability to catalyse the transamination of α-ketoglutarate **59** and phenylalanine **60** to glutamate **61** and 3-phenylpyruvate **62** (Scheme 12).

59 **60** **61** **62**

Scheme 12

Results showed that the overall catalytic efficiency of the conjugate with the lysine residue at position 51 was 4200-fold increased compared to unconjugated pyridoxamine phosphate. It also accelerated the reaction 12-fold when compared to the IFABP containing conjugated pyridoxamine, but without lysine mutations. The reactions with both conjugates were also shown to be stereoselective. To show that the increased catalytic activity occurred due to the introduction of the lysine residues, the two conjugates were incubated with an excess of α-ketoglutarate **59**. In the absence of an amino acid substrate, only half the catalytic cycle can be accomplished, resulting in the conversion of the pyridoxamine cofactor to the corresponding pyridoxal. In the presence of lysine, this should be able to form a Schiff base complex such as **54**. In both cases the UV absorbance spectra showed a band at 422-428nm, characteristic of a pyridoxal-Schiff base complex. In control experiments with conjugates without the lysine mutation, no such absorbance was observed, indicating that in the case of the lysine mutated conjugates the mechanism is *via* Schiff base formation as shown in Scheme 11.

3.2 The Selection Approach.

As described above, the traditional approach to artificial enzymes has been the rational design of active sites. This strategy has provided us with much information on the criteria required for successful catalysis, such as binding processes, positioning of functional groups, and the importance of the correct polarity within the microenvironment of the active site. The unfortunate reality is however, that from concept to catalysis is a very time-consuming process, and as shown by Deiderich's dendrophanes, the concept does not always work in practice as well as theory. To return to the drawing board, and adjust even small fragments of the design can take years. In an attempt to move away from this linear approach, several groups have started to use a selection strategy, which allows for the simultaneous screening of a wide range of possible catalysts thus significantly reducing the time required to detect the best hosts.

Traditionally, this has been the domain of biological chemistry as in the cases of catalytic antibodies and ribozymes, but the introduction of combinatorial chemistry, and improvements in screening techniques have opened up this strategy to the synthetic chemist.

The selection approaches can loosely be divided into two catagories: those which use binding of a transition state analogue as their selection criteria, and those which base their selection directly on catalytic activity. However, experimental practice has shown that whilst the initial concept is based on either one of these, the combination of both produces the most effective results.

3.2.1 Approaches Utilising Binding Capabilities as their Selection Criteria.

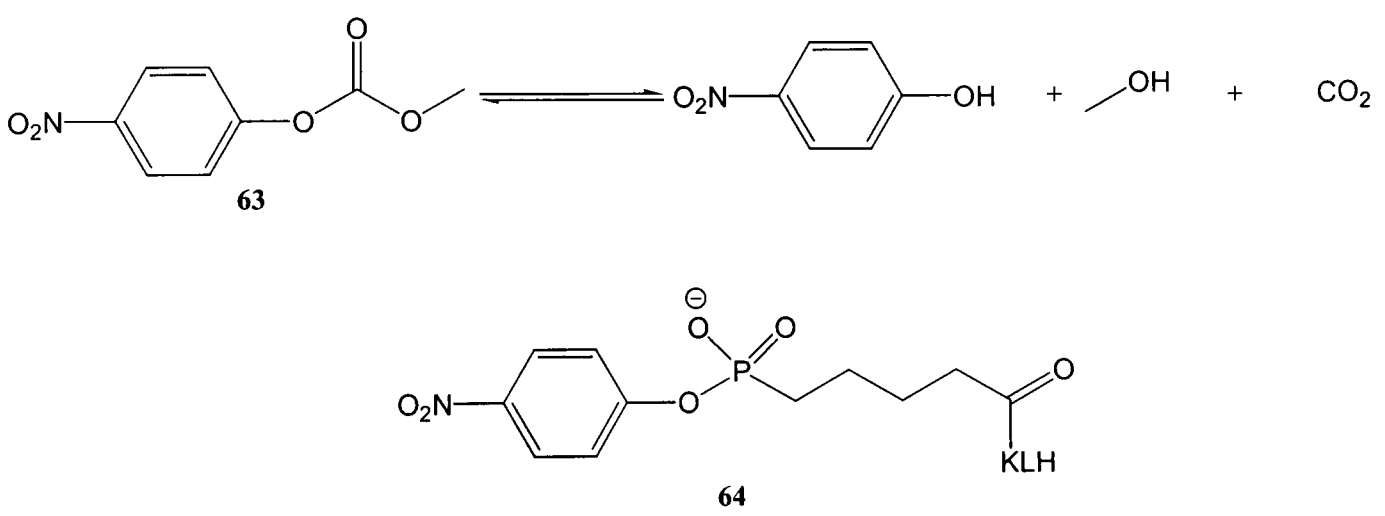
The earliest pioneers of selection approaches based their screening criteria on the ability of a host to bind the transition state analogue (TSA) of a given reaction. The logic behind this was that if a macromolecule showed an affinity for a molecule that resembled the transition state of a reaction both geometrically and electronically, then the host should act as an enzyme mimic for that reaction by preferentially binding, and stabilising the transition state of a reaction over the ground state.

3.2.1.1 Catalytic Antibodies.

Antibodies are proteins produced by the body in response to an antigen (alien species) which bind selectively to that molecule alone. As early as 1948, Pauling recognised that the fundamental processes that determine binding in enzymes and antibodies is the same[46]. In 1969, Jencks went one step further and surmised that an artificial enzyme could be prepared by raising antibodies against a molecule that is a stable representation of the transition state for the reaction of interest[47]. Raso and Stollar used a TSA to generate polyclonal antibodies that would exhibit catalytic activity for Schiff-base formation as part of a reaction analogous to pyridoxal-assisted amino acid metabolism[48]. Antibodies were formed that showed binding to the substrate, but no catalysis was observed. It was not until after the introduction of technology that allowed for the isolation of monoclonal antibodies[49] that Lerner[50] and Shultz[51] simultaneously reported the preparation and use of the first truly catalytic antibodies in 1986.

Early studies were purely based on a TSA approach. The TSA of a given reaction was designed and synthesised. As small molecules are not recognised by the immune system, the TSA had to be attached to a carrier protein, generally keyhole limpet hemocyanin (KLH) after which it is termed the hapten. The hapten was introduced into the bloodstream and immunisation allowed to occur. The desired monoclonal antibody was then selected from the polyclonal population based on its ability to bind the TSA.

Schultz exemplified this approach through his early work on the hydrolysis of carbonates (Scheme 13)[51]. The charge distribution and tetrahedral geometry of the intermediate formed in the hydrolysis of carbonate **63** was mimicked using the phosphonate ester **64** as a TSA. Antibodies raised against **64** exhibited both good substrate selectivity, and excellent rate accelerations of factors of up to 10^4. As a result of the lesser size of the *p*-nitrophenol product compared with the carbonate substrate, and due to the expulsion of carbon dioxide, catalytic turnover was also achieved.



Scheme 13

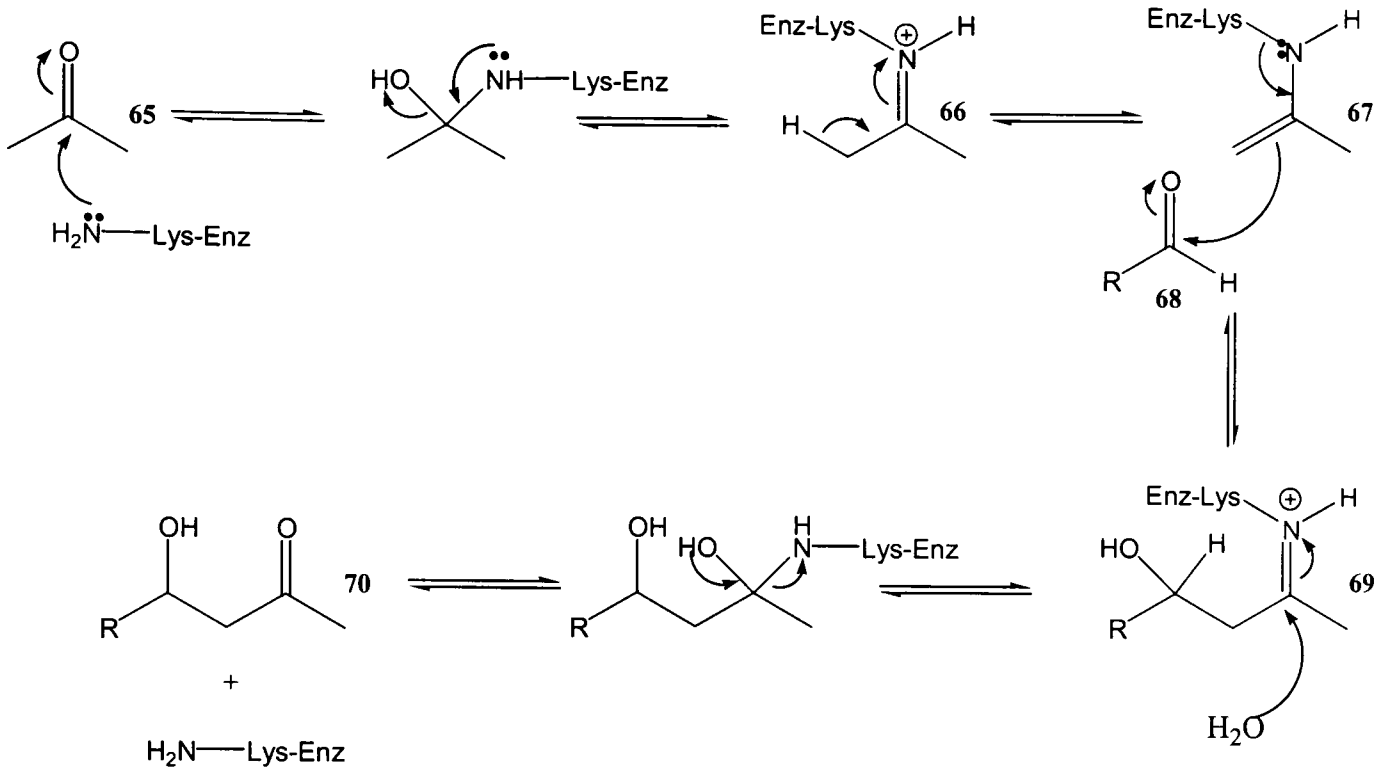
Using this primitive approach, catalytic antibodies have been generated for reactions including ester hydrolysis[50], transesterification[52], hydride reduction of carbonyl compounds[53], Diels-Alder cyclisations[54,55], ring-opening and closing reactions[56], sulfoxidations[57], carbohydrate deprotections[58], oxy-Cope rearrangements[59] and cationic cyclisations[60]. Although rate accelerations were observed in all cases, these did not come close to rivalling those observed in the case of natural enzymes. These disappointing results are the product of two flaws in the concept: inaccurate design of the TSA, and the selection event being based on TSA binding only. Significant

improvements have been made on both these fronts in recent years leading to some impressive results.

Thus, design of a TSA is problematic, since the transition state for a given reaction is not a real discrete entity, and in many cases there are a multiplicity of transition states to consider. Furthermore, any TSA can only be considered as an approximation to the true charge distribution required. To compensate for this, recent developments have been to design a hapten that not only mimics the transition state, but also has the role of placing functional groups at decisive positions within the antibody-binding site. There are three main strategies here: reactive immunisation[61,62], bait-and-switch haptens[63], and heterologous immunisation[64].

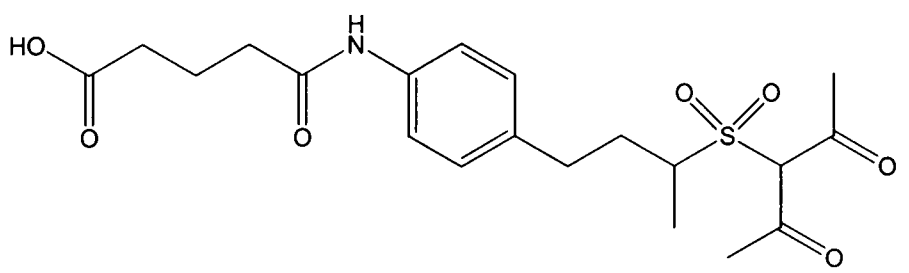
In reactive immunisation a reactive antigen is designed so that a chemical reaction or reactions, such as the formation of a covalent bond, occurs in the binding pocket of the antibody during immunisation. The chemical reactivity programmed into the antibody is designed to be an integral part of the reaction when the corresponding substrates are used. Lerner *et al.* have utilised this approach in the design of aldolase catalytic antibodies[65].

A lysine residue in the active site of natural aldolases is crucial to the mode of catalytic action (Scheme 14)[66]. Attack of this lysine on the substrate **65** forms a complex **66** in which the enzyme is covalently bound to the substrate *via* a Schiff base linkage. Tautomerisation of this imine **66** forms an enamine **67** that reacts with the second carbonyl substrate **68** to form another Schiff base **69**. Hydrolysis with water affords the aldol product **70** and releases the active lysine group.



Scheme 14

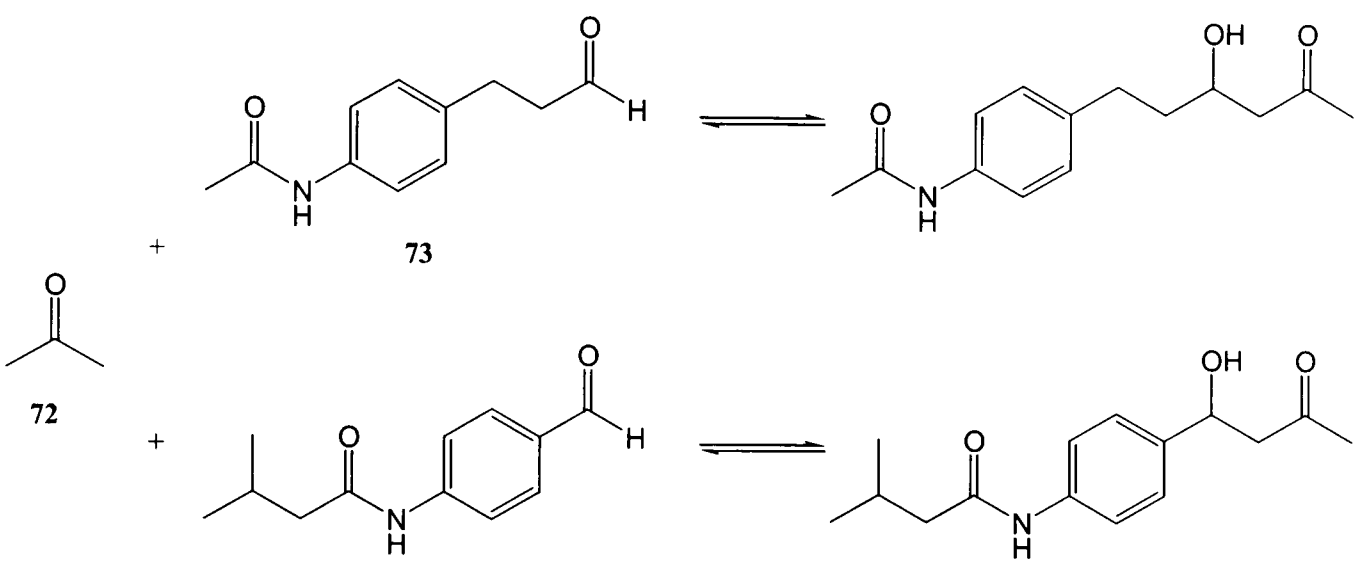
Lerner used the β-diketo sulfone hapten **71** to generate antibodies capable of catalysing an aldol reaction (Figure 14)[65]. The tetrahedral geometry found in the transition state of the C-C bond forming step is mimicked using the sulfone moiety. The role of the β-diketone moiety is to introduce a lysine residue into the active site of the antibody. Lerner reasoned that if an antibody had an appropriately placed lysine residue it would attack one of the hapten carbonyl groups to form a carbinolamine that should collapse to a Schiff base. By virtue of the acidity of the central proton of the β-dicarbonyl unit, this can tautomerise to form an enamine, which due to a second carbonyl functionality in the β-position is a stable vinylogous amide. This has a strong UV absorption outside the range of the protein (316nm), and thus instead of screening for binding, catalysts were selected by screening for the new absorption.



71

Figure 14

Seventeen monoclonal antibodies were isolated in this way. Before testing for catalysis, the antibodies were first screened for their ability to covalently react with 2,4-pentanedione to form a stable enaminone. Nine of the seventeen antibodies showed the enaminone UV absorption at 316nm after incubation with 2,4-pentanedione. All nine were then found to catalyse the aldol reactions between acetone **72** and 3-(4-acetamidophenyl) propanal **73** and 4-isobutyramidobenzyaldehyde **74** with rate accelerations greater than 10^5 fold (Scheme 15). The same antibodies also showed activity in catalysing retro-aldol reactions. All the reactions followed Michaelis-Menton kinetics, and were inhibited by 2,4-pentanedione, which confirmed that catalysis was occurring *via* the proposed covalent mechanism.



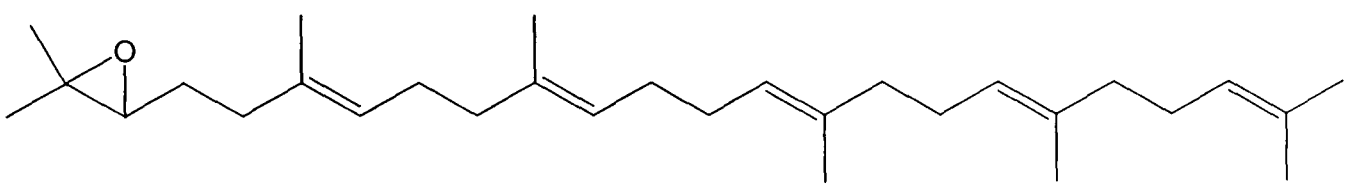
Scheme 15

The kinetic capabilities of these antibodies for a variety of substrates were compared with the commercially available aldolase antibody 38C2. In all cases the catalytic proficiency was found to be superior to the commercially available antibody, and in the particular case of the antibody named 84G3, activity was found to be 1000-fold higher than for any other reported catalytic antibody[67]. Using reactive immunisation clearly shows great potential, and this example represents the most efficient antibody catalysts prepared to date.

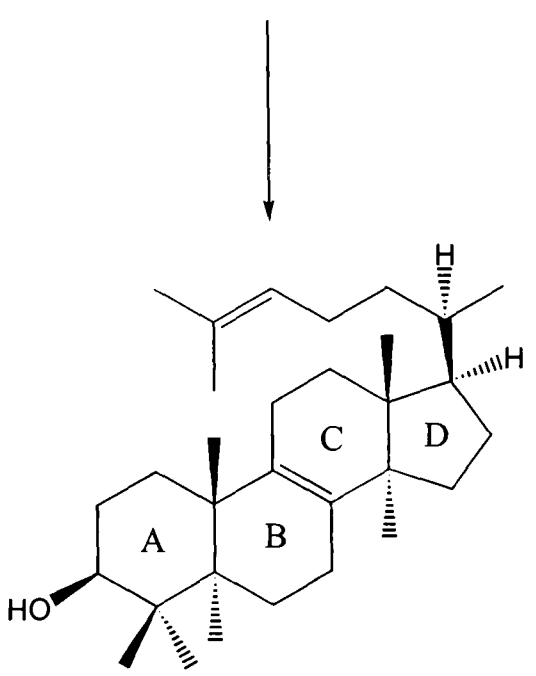
In the "bait-and-switch" approach, haptens are designed which contain strongly charged or polarised functional groups aimed to act as a "bait" to the immune system[63]. Counter-charged or polarised residues are thus obtained in the binding site. When we "switch"

from the inducing antigen to the substrate, these amino-acid residues are expected to play an important role in catalysis.

Terpenoid cyclase catalysed carbocation cyclisation cascades are some of the most important and complex carbon-carbon bond forming reactions in chemistry and biology[68]. For example, in the biosynthesis of cholesterol, the cyclisation of the open-chain polyene squalene oxide **75** to yield the tetracyclic product lanosterol **76** is perfomed in a single chemical reaction catalysed by lanosterol synthase (Scheme 16). Seven stereocentres are formed during the cascade reaction, and remarkably out of a possible 128 stereoisomers, only one is formed.



75



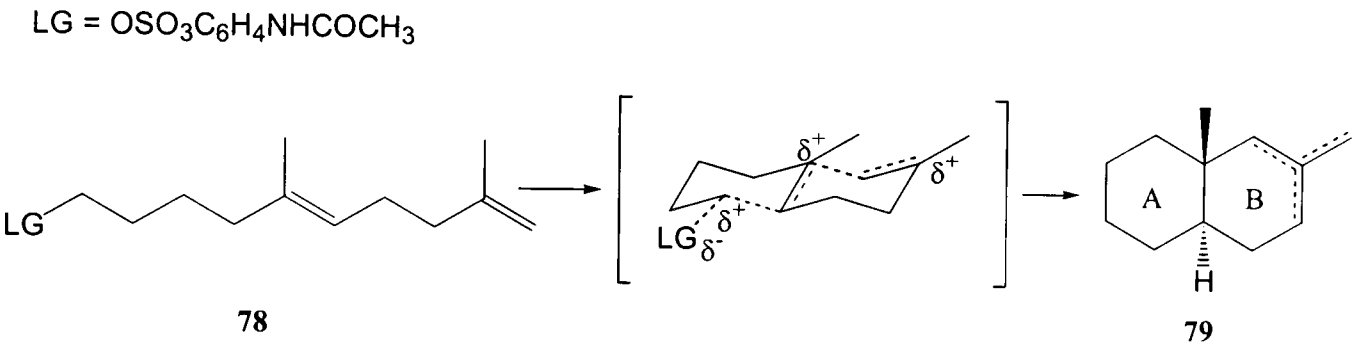
76

Scheme 16

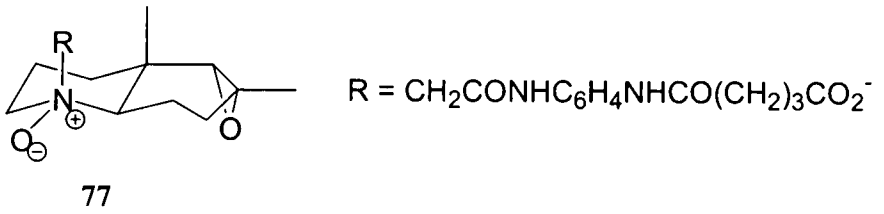
The active site of terpene cyclases is nested deep within an α-helical superstructure, where numerous hydrophobic residues have two roles: to sequester the linear polyene substrate from solvent, and to form a template that forces the substrate into the correct conformation and stereochemistry required for cyclisation. Aromatic residues within the active site are positioned to stabilise any carbocation cyclisation intermediates through

cation π-interactions. Polar residues in the active site participate in electrostatic stabilisation and govern the regiochemistry of deprotonation and protonation steps.

Lerner recognised that development of a catalytic antibody which could imitate terpene cyclases was a desirable target, since it would be possible to engineer the catalyst to maximise the diversity of cyclisation products, and also enable the use of unnaturally occurring polyene substrates[69]. He designed hapten **77** to mimic the productive chair-chair conformation required for the cyclisation of polyene **78** to the *trans*-decalin **79** (Scheme 17).



78 **79**



77

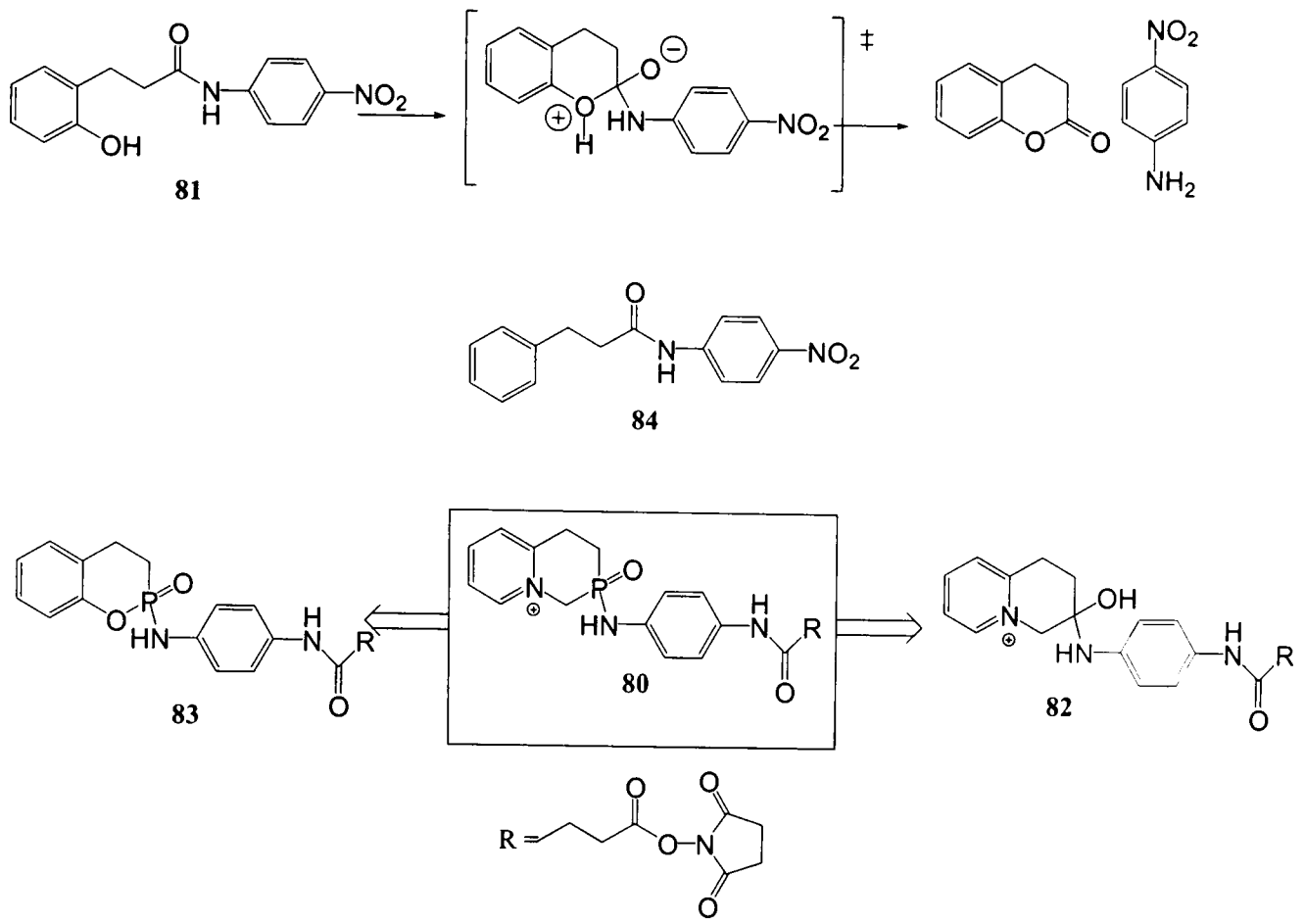
Scheme 17

The "bait and switch" strategy was incorporated into the design by the introduction of the *N*-oxide moiety, which was expected to elicit polar functional groups in the antibody active site that would initiate the cyclisation reaction by facilitating the departure of the leaving group. The epoxide moiety was introduced into the hapten in the hope that it would encourage termination either by fixing the B ring of the hapten in the half-chair conformation that could favour proton elimination, or by eliciting a polar residue that could interact with a water molecule that could in turn act as a nucleophile for attack on the final carbocation.

Use of this hapten generated an antibody HA5-19A4 that, when reacted with **78**, produced the desired *trans*-decalin **79**. X-ray crystallography revealed that, as designed, the *N*-oxide moiety had elicited an aspargine residue in a position to stabilise the developing negative charge on the aryl sulfonate leaving group. The rest of the active site was lined with hydrophobic aromatic side chains. The epoxide was found not to have elicited any polar residue that could encourage nucleophilic formation of a tertiary alcohol, and instead elimination occurred to form three regioisomeric "decalenes" (Scheme 17). Recently the "bait-and-switch" strategy has also been successful in developing antibodies that trigger and control multi-ring cyclisations starting from an epoxide moiety as seen in sterol biosynthesis[70].

Heterologous immunisation consecutively uses two related hapten structures in the immunisation of a single animal[64]. By carrying different functional groups, the two haptens can stimulate stereoelectronic features of the transition state that cannot be readily incorporated into a single hapten. Ersoy *et al.* used this approach to investigate the generation of antibodies capable of amide hydrolysis[64].

Ersoy envisaged that hapten **80** should be a good TSA for the amide hydrolysis of substrate **81**[64]. He recognised that the acyl transfer step in serine protease catalysed reactions was crucial, and had made several futile attempts to incorporate a nucleophilic residue in the active site of the antibody. As a different approach, he incorporated a phenolic nucleophile into the substrate **81**. The functional groups on hapten **80** were separated and two haptens, **82** and **83** synthesised (Scheme 18). The pyridinium salt on **83** was intended to induce a basic residue in the antibody to accept the proton from the attacking phenolic hydroxyl group. This protonated residue could then deliver its proton to the amine functionality of the tetrahedral intermediate formed in the reaction, and thus facilitate its departure. The tertiary alcohol was to serve as an isosteric replacement for the tetrahedral intermediate. The phosphonamidate hapten **82** was designed to induce an acidic residue in the active site to stabilise the developing oxyanion of the transition state. The six-membered ring containing the phenol and phosphonamidate was intended to ensure the correct orientation of the hydroxyl group such that nucleophilic attack on the amide could occur. Antibodies were then generated against each of the haptens, and heterologously by sequential immunisation of both haptens.



Scheme 18

Hydrolysis of substrate **81** in the absence of any antibodies was found to be very slow, with a rate of k=4.7 x 10^{-7}min^{-1}, and almost identical to that of substrate **84** without the phenol group. In the presence of antibodies generated using only one of the two haptens, catalysis was observed, with a k_{cat} of 1.8 x 10^{-3}min^{-1}. The antibodies generated using the heterologous immunisation method were however over 10 times more effective than those antibodies generated homologously, and catalysed with k_{cat} of 1.5 x 10^{-2}min^{-1}. Interestingly, none of the antibodies catalysed the hydrolysis of amide **84**, which was a clear indication that the phenolic moiety was essential for catalysis to occur.

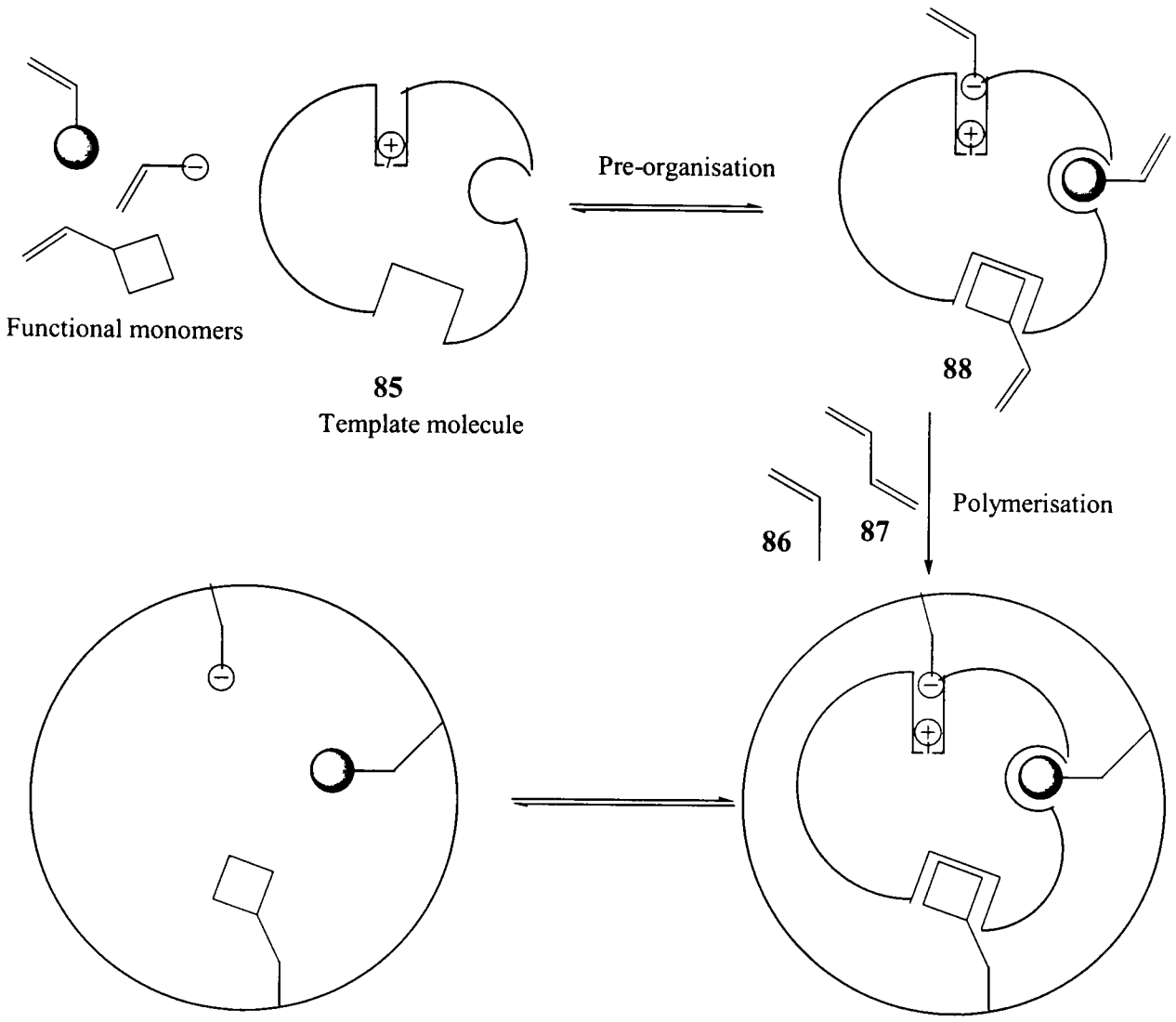
Perhaps even more importantly the other problem of selection has also been addressed. Selection techniques have been shifted in the direction of catalysis over affinity. As antibody library sizes increase, more sensitivity is required in screening techniques. Advances in technology have meant that this can be achieved through the use of chromogenic or fluorogenic assays, ESI mass spectrometry, or through the selective amplification of catalytic events[62].

As shown in the examples above, the field of catalytic antibodies has made significant advances since Schultz and Lerner first reported their achievements in 1986. There are still many problems associated with this technique, despite these impressive successes. The antibody is generated by the immune system only for hapten affinity, and thus only develops the tools of molecular recognition necessary to obtain the required dissociation constant (10^{-4} to 10^{-10}M). It may therefore be unnecessary for the immune system to recognise all the features of hapten design. The sacrifice of mice for the generation of antibodies is not desirable, although introduction of *in vitro* immunisation techniques has made this less of an issue. The major drawback however, is that the biological techniques involved in generating and isolating catalytic antibodies are both complex and time consuming, such that the whole process from design of the hapten, to structural characterisation of the active site can take years.

3.2.1.2 Molecularly Imprinted Polymers.

Whilst antibodies have shown promising results, they share many problems with enzymes such as lack of thermal stability and chemical robustness which make them inappropriate for use on a large scale. A synthetic analogue would therefore be of much interest. With this in mind, molecular sized cavities have been generated in the solid state by polymerisation in the presence of a template molecule. If the template is a TSA then the polymer formed should behave as an artificial enzyme towards the reaction chosen. This technique, first introduced by Wulff, is known as molecular imprinting[71].

The polymerisation process is shown in Scheme 19. The first step is the pre-organisation of functional monomers around the template molecule **85**. The functional groups are chosen such that they should interact with specific binding sites on the template. Radical polymerisation of a monomer **86**/cross-linker **87** mixture around the pre-organised complex **88** forms a macroporous polymer **89** in which the template molecule is bound at specific binding sites. Removal of the template molecule from the polymer leaves cavities of the correct size and shape of the template, and bearing appropriately placed functional groups for binding and catalysis.



Scheme 19

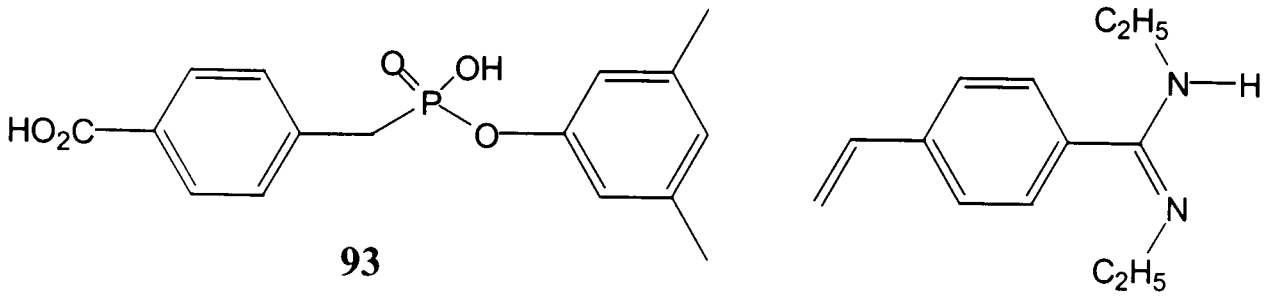
Two methods are used to appropriately position the functional monomer groups[72]. The first of these is the self-assembly approach, in which electrostatic interactions are used to bind together the template and the functional group. In this instance the pre-organisation step is actually a dynamic equilibrium, with the free and complexed forms in constant exchange. As a result, the molecular recognition sites formed within the polymer are heterogeneous, which can lead to problems in the catalytic applications of these polymers. The second method is the controlled distance approach, in which the functional monomer is actually a conjugate of the template.

The self-assembly approach has proved to be the more popular, since the potential for a "better fit" is greater if the intermolecular bonds are strong enough. It is thought that as hydrogen bonds and Van der Waals forces are strongly dependant on bond lengths, the controlled method will not necessarily place interacting groups the correct distances

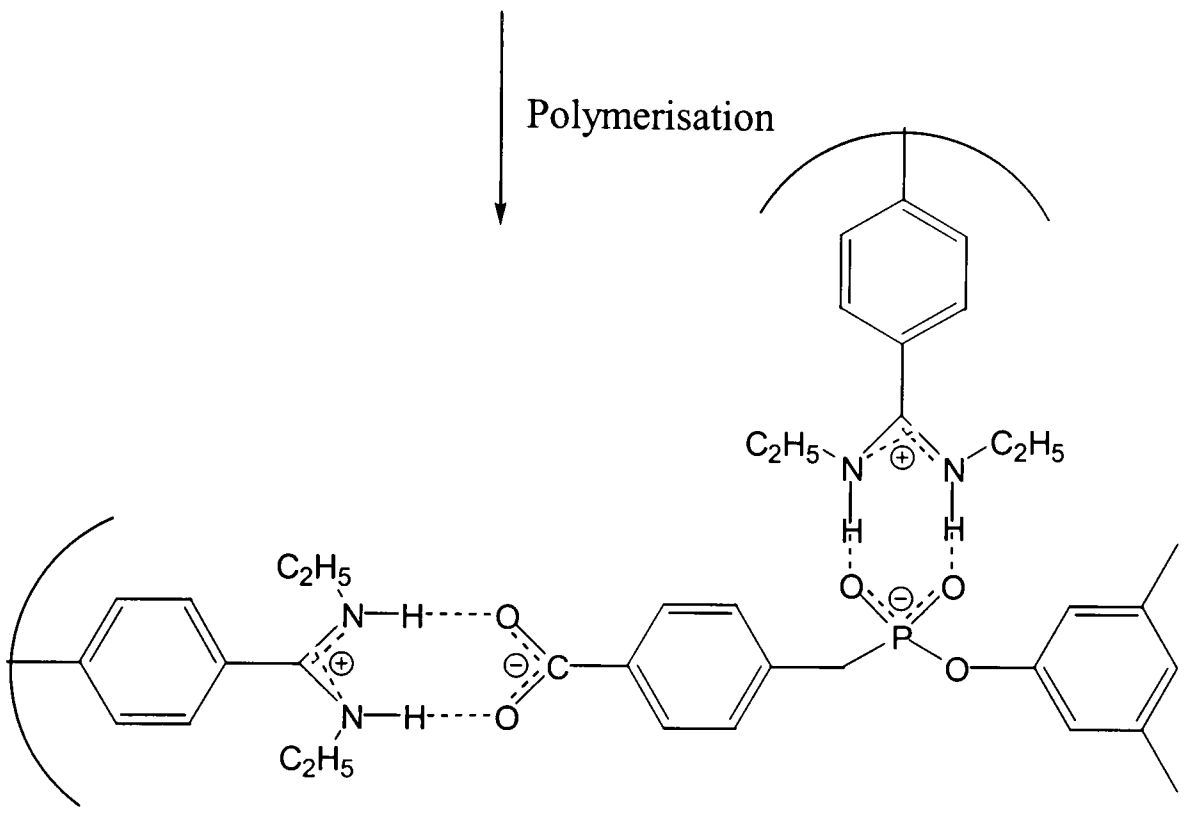
apart. Molecular imprinting has been used with great success to produce chiral stationary phases for chromatography and for selective extractions, and the goal now is to apply this methodology to finding efficient enzyme like catalysts.

Molecularly imprinted polymers formed using a TSA have shown modest success in being applied as enzyme catalysts. Rate enhancements have been observed, most notably in ester hydrolysis[73], aldol reactions[74] and Diels-Alder cycloadditions[75], but these are generally in the value of between 1.6 and 100-fold, and do not approach the activity of enzymes. However, as with catalytic antibodies, researchers have realised that recognition of a TSA alone is not enough to confer catalytic activity.

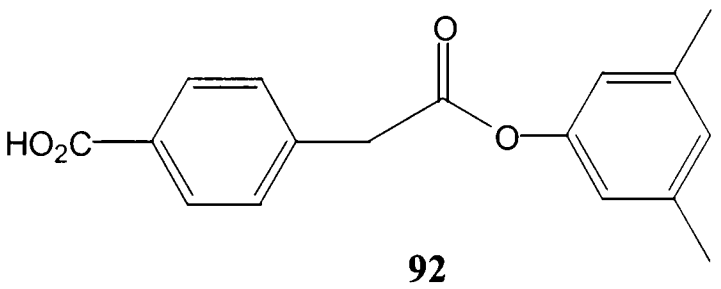
Arginine residues play a crucial role in many natural enzymes, as the guanidinium group is excellently suited to binding oxyanions such as carboxylates and phosphates. In recognition of this fact Wulff designed a molecularly imprinted polymer **91** containing an appropriately positioned amidine group, to catalyse the hydrolysis of ester **92**[73]. The thinking was not only that this should stabilise the transition state of the reaction, but also polarise the carbonyl moiety of the ester group to make it more receptive to attack from incoming nucleophiles. The polymer was imprinted with TSA **93**, a phosphonate chosen for its tetrahedral geometry mimicking that found in the intermediate formed during ester hydrolysis (Scheme 20).



93



Molecularly Imprinted Polymer **91**

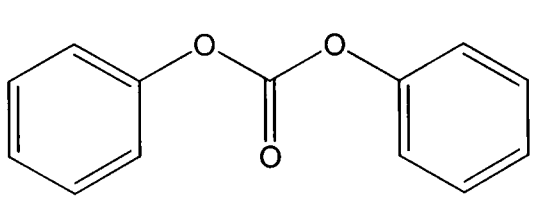


92

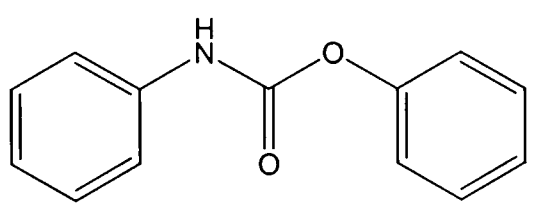
Scheme 20

The molecularly imprinted polymers formed after imprinting with **93** were found to cause a 100-fold acceleration of the hydrolysis of ester **92**, and also to follow Michaelis-Menton kinetics.

Wulff used a similar amidine-phosphate complex to form molecularly imprinted polymers capable of catalysing the hydrolysis of diaryl carbonate **94** and aryl carbamate **95** (Figure 15) using supension polymerisation techniques[76]. Carbonates were chosen since they do not form stable anionic products which in ester hydrolysis are thought to act as reaction inhibitors. The polymers formed catalysed hydrolysis of **94** and **95** by a factor of 588 and 1435 respectively. Interestingly, the polymer which was randomly functionalised with the amidine monomer, was also found to be a significantly active catalyst. The rate enhancements shown by the molecularly imprinted polymer were only 6-24 fold when compared to this reference polymer. In contrast the amidine functional monomer in solution showed negligible catalysis at the reaction pH. This is by far the largest rate accleration for ester hydrolysis achieved thus far with molecular imprinting techniques, but the results clearly show that non-specific catalysis at the surface of the polymers also has a role to play in rate enhancements.



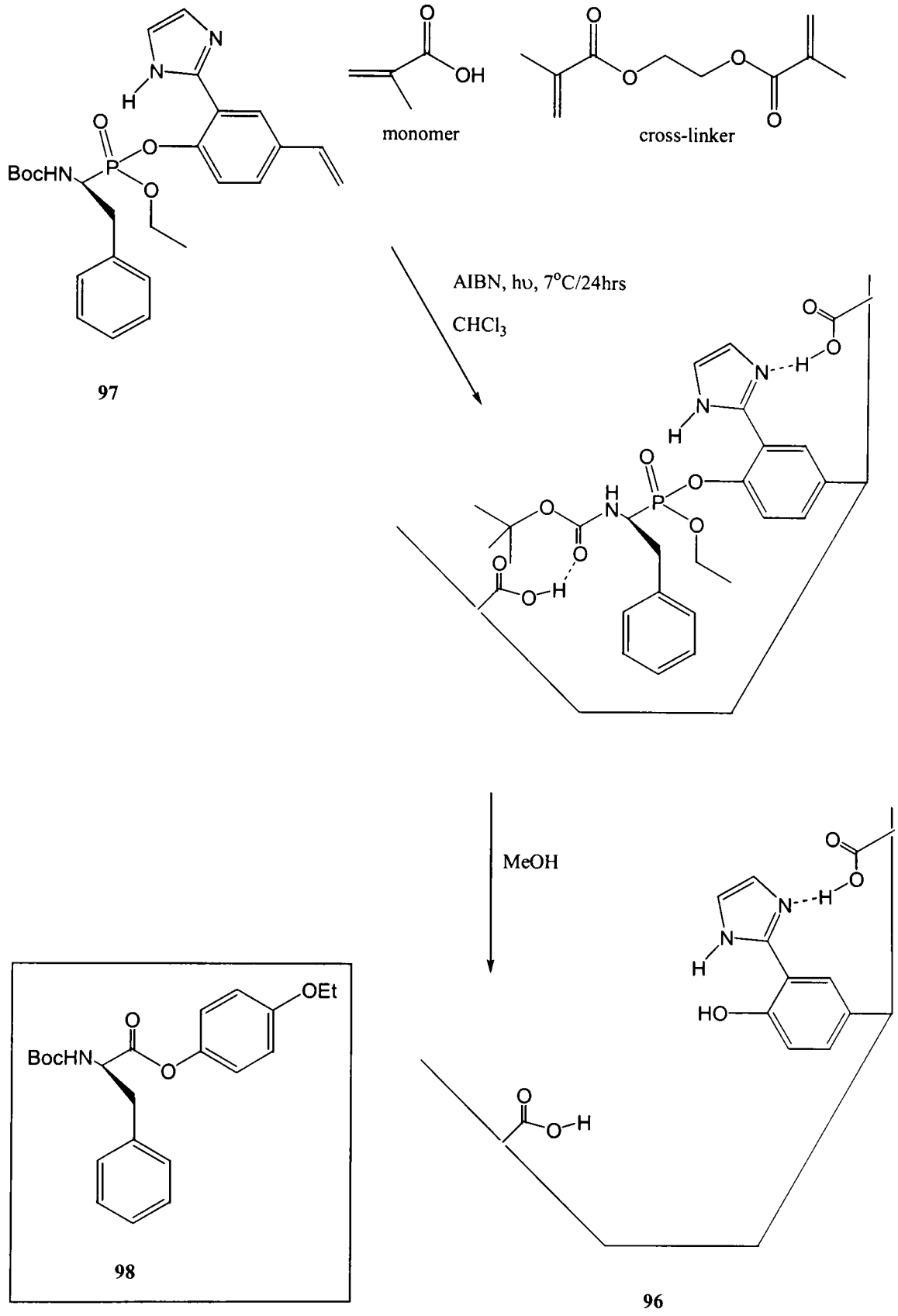
94



95

Figure 15

By combining both noncovalent and covalent binding forces between the template and functionalised monomers Shea has synthesised molecularly imprinted polymers that incorporate the elements believed to be responsible for the catalytic action of chymotrypsin[77]. The polymer **96** imprinted with template **97** should have, after removal of the template molecule, an active site complementary to the D enantiomer of BocPheOEt **98**. A phenolimidazole catalytic group should be positioned in close proximity to the carbonyl group of the bound substrate. Hydrogen bonds between carboxylic acid groups in the polymer, and complementary hydrogen bonding groups on the substrate should enable selective binding, and the tetrahedral geometry provided by the phosphonate group should create a site complementary to a transition state structure. It was hoped that an extra carboxylic acid group should also be hydrogen bonded to the imidazole group, and thus complete a "catalytic triad" similar to that observed in chymotrypsin (Scheme 21).



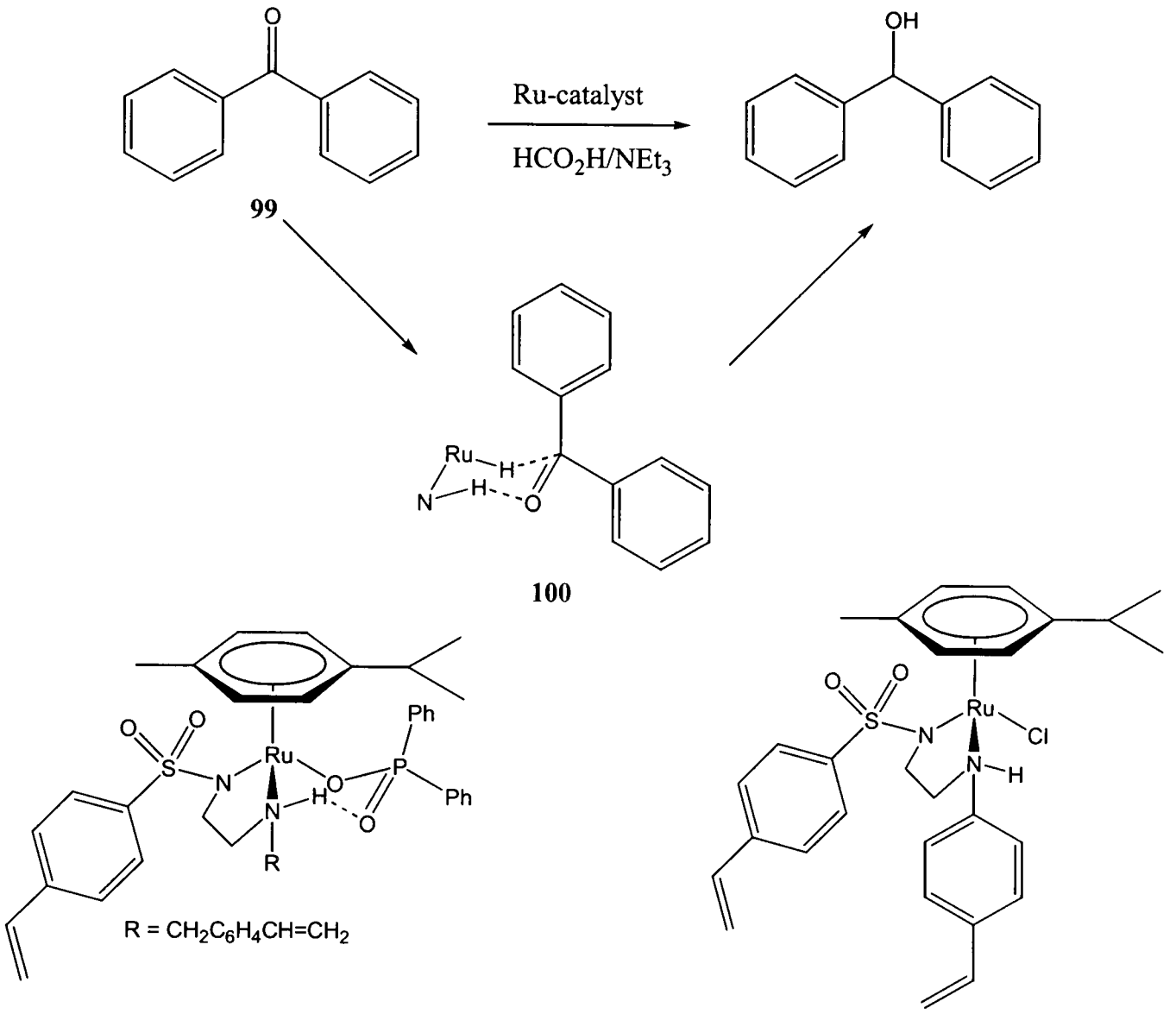
Scheme 21

To investigate which functional groups were key elements to the catalysis, a variety of polymers were also designed without the imidazole group, with random placement of the carboxylate groups, without the tetrahedral complementarity, and with achiral templates such that no stereospecificity should be observed. When BocPheOEt was hydrolysed in the presence of polymer **96**, complete stereoselectivity for the D-isomer was observed, and catalysis occurred with 10-fold rate acceleration. The catalyst also

exhibited turnover, performing 5 catalytic cycles. Use of the control polymers showed less significant catalysis, and revealed that all the catalytic elements incorporated were required. Although the rate accelerations obtained were relatively modest, the 100% enantioselectivity observed must be considered a significant achievement.

The synthesis of artificial metalloenzymes is a highly desirable target. The activity and selectivity of metalloenzymes is determined decisively by the ligands of the second coordination sphere, i.e. by the microenvironment of the catalytically active centre. Molecular imprinting provides a method to control the microenvironment of synthetic transition metal catalysts. In this approach, a catalytically active metal substrate with a spectator ligand containing one or more polymerisable side chains is coordinated to a pseudo substrate. The complex should represent the transition state of the catalytic transformation. Polymerisation in the usual manner, and decomplexation of the psuedo substrate should yield an immobilised catalyst with a specific cavity in close proximity to the catalytically active metal centre. Mosbach prepared an aldolase-mimicking synthetic polymer in this manner, but the metal complexes were prepared *in situ* prior to the polymerisation, and not characterised[74].

More recently Severin has reported the synthesis of imprinted ruthenium catalysts that accelerate the transfer hydrogenation reaction of benzophenone **99**[78,79]. Half sandwich complexes of ruthenium containing amine-based ligands are known to catalyse this reaction with high rate accelerations and enantioselectivity. The reaction proceeds *via* a six-membered transition structure **100**. Severin prepared a polymer using the fully characterised phosphinato ruthenium complex **101** as a TSA (Scheme 22)[79]. The role of the diphenylphosphinato ligand was to act as a pseudo substrate for benzophenone **99**. A control polymer was also synthesised, using the ruthenium complex **102**. This polymer should contain the same amount of active ruthenium, but without the specific microenvironment surrounding it as a result of the diphenylphosphinato ligand.



Scheme 22

The transfer hydrogenation of benzophenone **99** was catalysed by both polymers, but the one that was imprinted with the TSA **101** was more active than the control polymer by a factor greater than five. Both polymers also exhibited turnover. This is quite a remarkable rate acceleration when we consider that both the imprinted and control polymers contain the same percentage of ruthenium complex. The substrate specificity of the two polymers was also examined using substrates **103-109** (Figure 16). Competition experiments using equal amounts of benzophenone **99** and a second ketone in the presence of the imprinted polymer always showed a preference for benzophenone as the substrate. In the control polymer, no significant preference for either substrate was observed. Severin has therefore successfully prepared a highly active, substrate specific artificial metalloenzyme using the molecular imprinting technique.

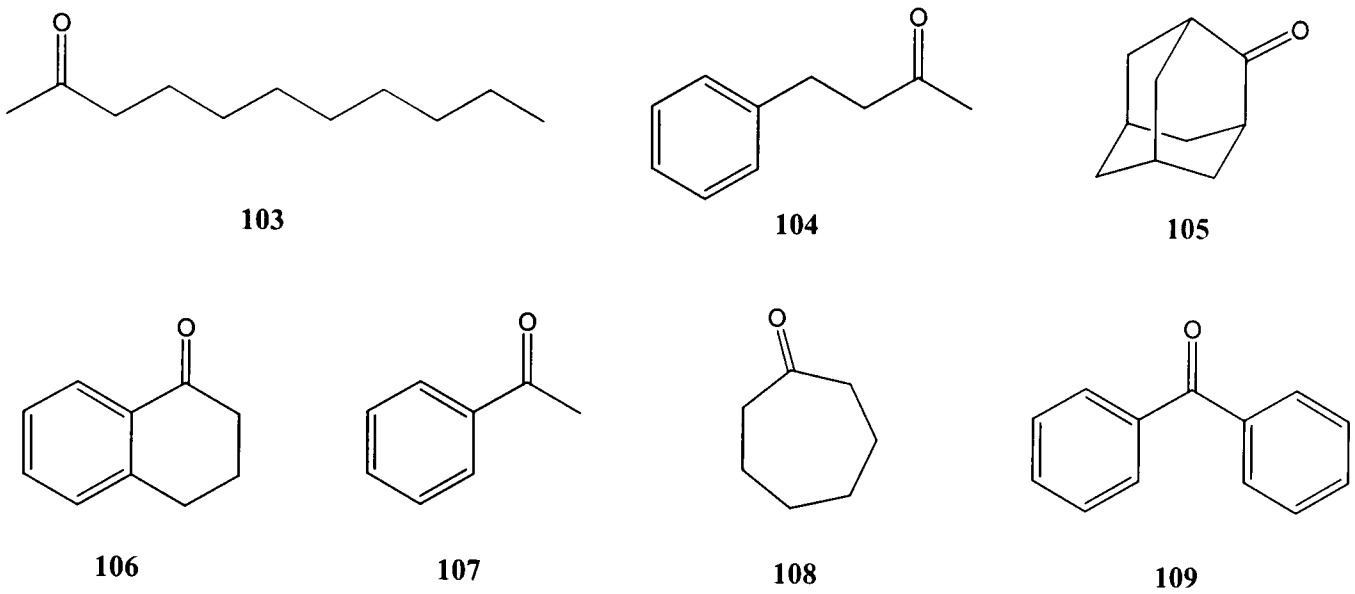


Figure 16

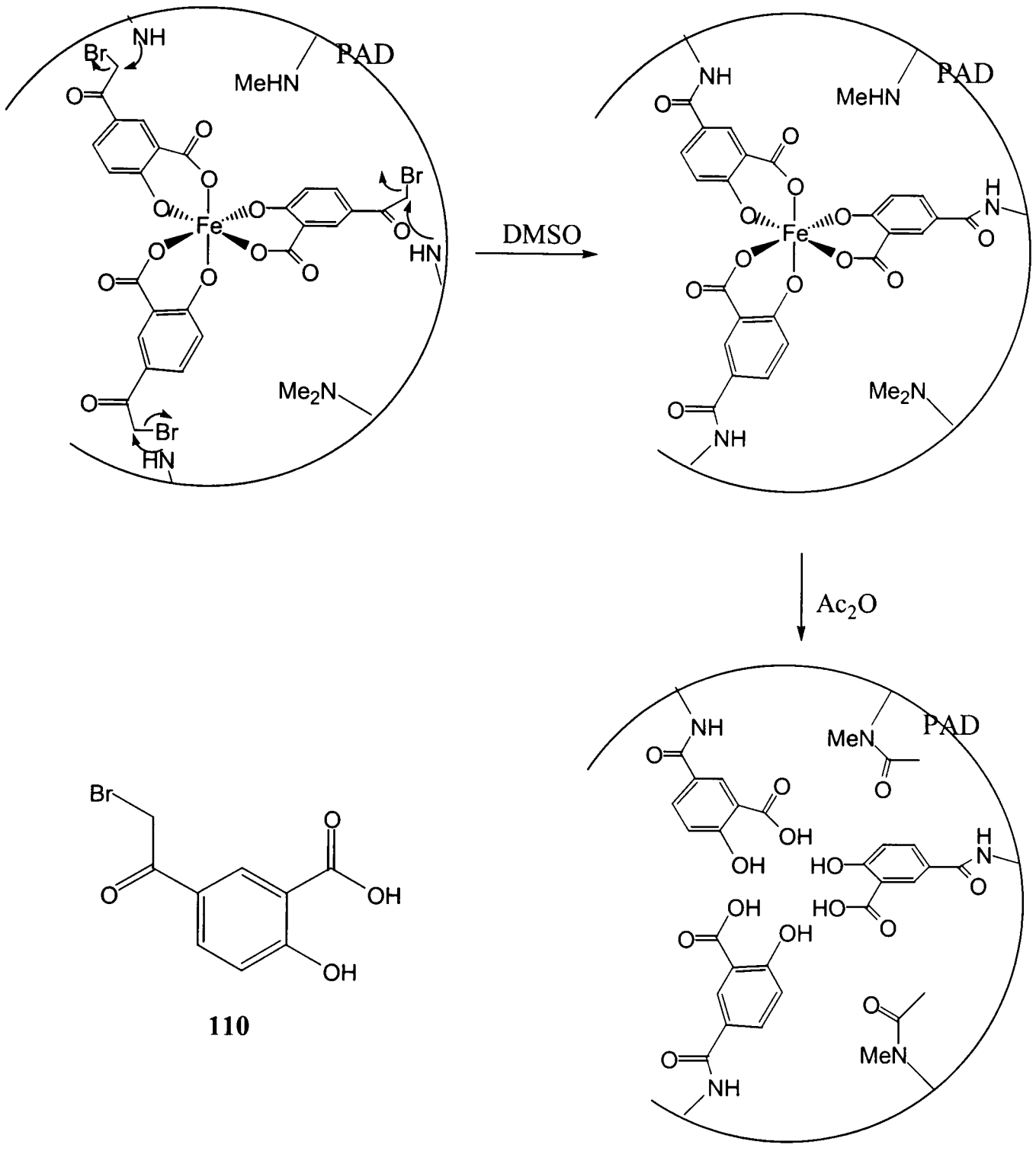
Although the concept of using molecularly imprinted polymers as artificial enzymes is very elegant, the results obtained so far using this technique have been relatively modest. One of the major problems of imprinted polymers is lack of homogeneity provided by the active sites, since polymerisation itself is a kinetic process. Another significant feature of enzyme catalysis is the ability of the enzyme to change it's conformation to welcome the substrate, and expel the product. Molecularly imprinted polymers are, by comparison, very rigid structures, and as such, product inhibition is often a major problem.

3.2.1.3 Imprinting an artificial proteinase.

J. Suh proposed that if a scaffold is used to assemble functional groups, and the pre-assembled functional groups are then cross-linked with a synthetic macromolecular spacer, then the removal of the scaffold would produce an active site with functional groups in close proximity to each other[80]. This idea is very similar to that of molecular imprinting, but utilises a pre-formed polymer backbone as the cross-linker. Using this method Suh intially developed a soluble polymeric artificial proteinase whose active site contained three hydroxyl groups[80].

More recently, Suh has developed this methodology to generate an insoluble aspartic protease analogue[81]. He used an Fe(III) ion as his scaffold, and preorganised three molecules of 5-(bromoacetyl)salicylic acid **110** around it. The complex was attached to a poly(aminomethylstyrene-co-divinylbenzene) (PAD) backbone, and the Fe(III) ion

removed to yield the polymer **111** with three salicate groups in close proximity (Scheme 23).



110 **111**

Scheme 23

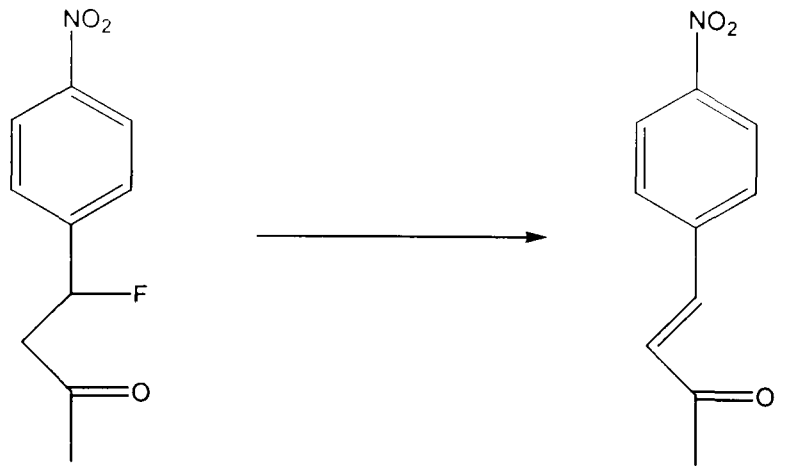
The trisalicylate units in **111** contain three carboxyl groups and three phenol groups as well as the amino groups of the PAD backbone. The carboxylate groups might be expected to mimic the side chain of aspartic acid, and the phenol groups are similar to the tyrosine side chain. The polymer acted as an aspartic protease analogue in the cleavage of bovine serum albumin (M_r 66,000) at a variety of pH values, but found it's optimum value (k_0=3.1 x 10^{-3}min^{-1}) at pH3, the same optimum value as for natural

aspartic proteases. The activity of **111** was much higher than for a randomly functionalised PAD, in which the salicylate residues are positioned some distance apart from each other. This indicated that the catalysis was occurring through participation of two or more salicylate residues. Catalysis could in theory either be due to colloboration between two carboxylate groups, or between a phenol and carboxylate group. That the optimum pH value for catalysis is 3, indicates that it is the former, as in natural aspartic proteases, since pH6-8 would be the optimum value for a phenol-carboxylate collaboration.

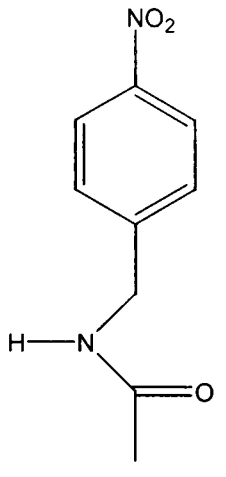
3.2.1.4 Bioimprinting.

It has recently been shown that the specificity and activity of enzymes and proteins can be modified when used in anhydrous or nearly anhydrous media[82,83]. New binding sites can be introduced into proteins using the bioimprinting technique, in which ligands are used to alter protein conformations in organic solvents[84]. This technique has now been adapted to generate artificial enzymes.

Slade and Vulfson induced catalytic activity in proteins by lyophilisation in the presence of a TSA[85]. The β-elimination of **112** (Scheme 24) has been studied extensively using catalytic antibodies and molecular imprinting. β-Lactoglobulin was imprinted by freeze-drying in the presence of TSA **113**. The sample produced was dried, and its ability to catalyse the β-elimination of **112** in anhydrous acetonitrile at 50°C investigated, and then compared to that of a non-imprinted control prepared under the same conditions. The imprinted protein showed catalytic activity more than three times that of the non-imprinted control, and 4 orders of magnitude higher than spontaneous β-elimination. This result seems relatively modest when compared to some of the rate accelerations obtained using catalytic antibodies, but was actually found to be an almost identical rate acceleration to that observed using molecularly imprinted polymers.



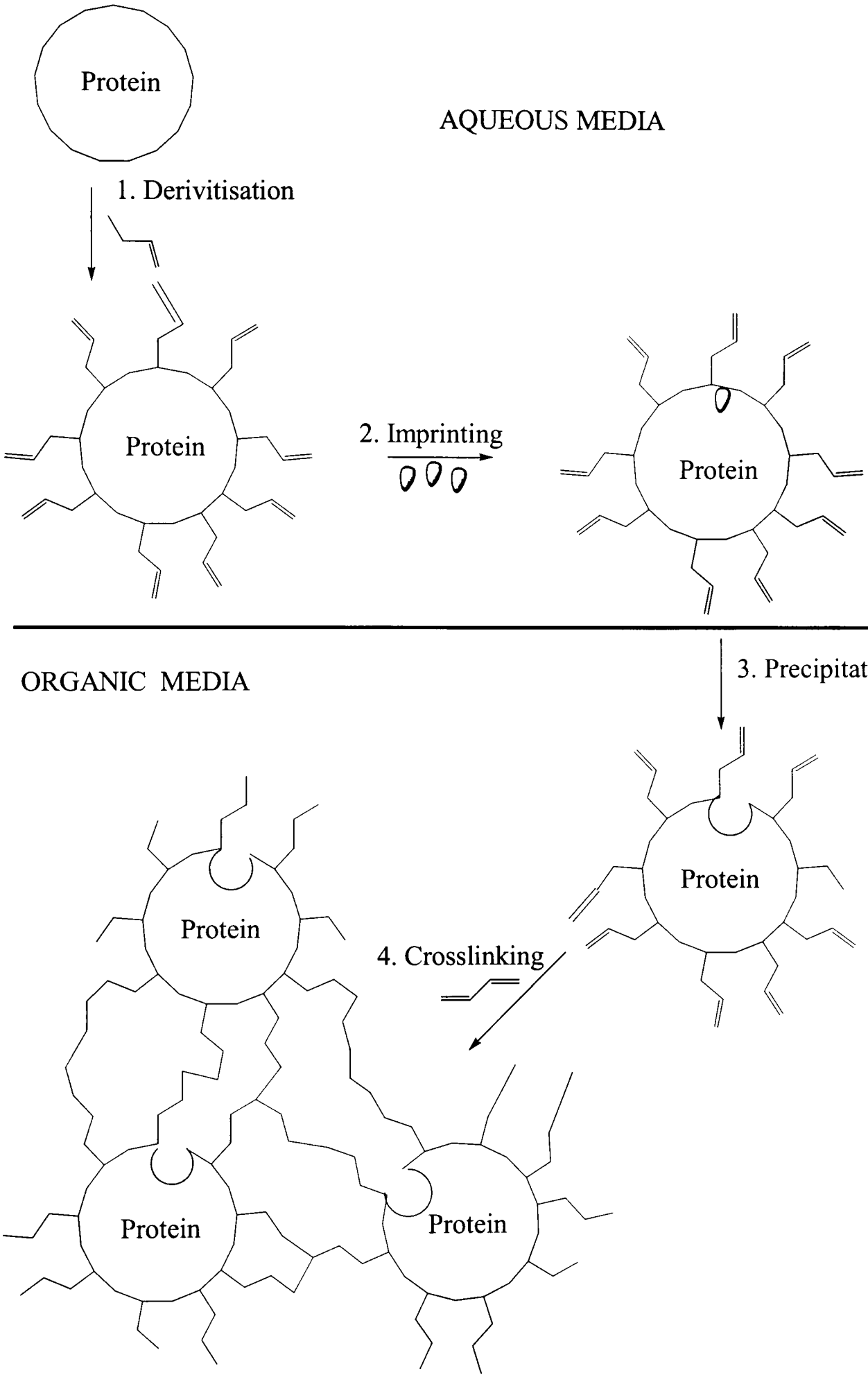
112



113

Scheme 24

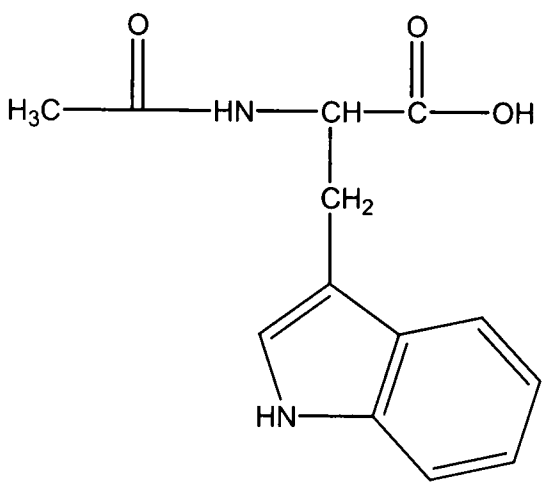
A major drawback of this method of imprinting is that the enhanced properties of these proteins are only available in nearly anhydrous environments, since an aqueous environment causes renaturation of the protein and consequential loss of any imprinting effect. Peifβker and Fischer reported a novel way to avoid this problem, by combining imprinting with a two step immobilisation method, with the result that the imprinted enzymes permanently retain their new structure in aqueous media (Scheme 25)[86].



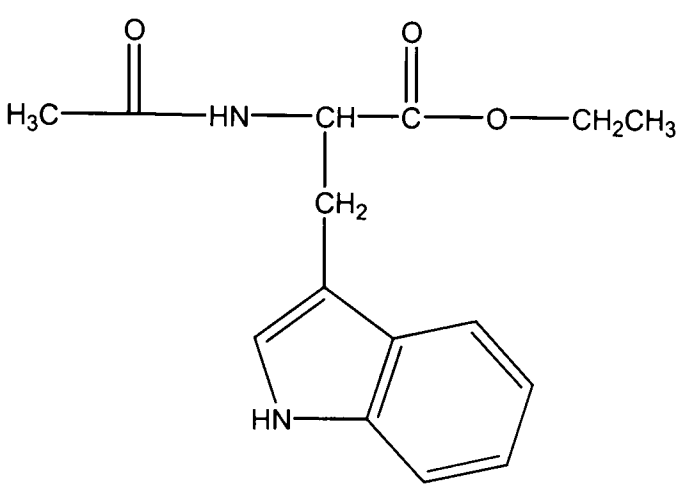
Crosslinking maintains the manipulation caused by imprinting

Scheme 25

Two enzymes, α-chymotrypsin, and subtilisin Carlsberg were derivitised using itaconic anhydride, to introduce polymerisable vinyl groups onto the protein. The itaconic anhydride functionalised the amino groups of lysine, hydroxyl groups of tyrosine, and thiol groups of cysteine. The derivatised proteins, and also non-derivatised proteins were imprinted with N-Ac-D-trp **114** (Figure 17) in buffer solution, precipitated using propanol, and dried. They were then either tested directly for their ability to hydrolyse the ester N-Ac-D-trp ethyl ester **115** in buffer solution, or crosslinked with ethylene glycol dimethacrylate using radical polymerisation, and then tested for catalysis.



114



115

Figure 17

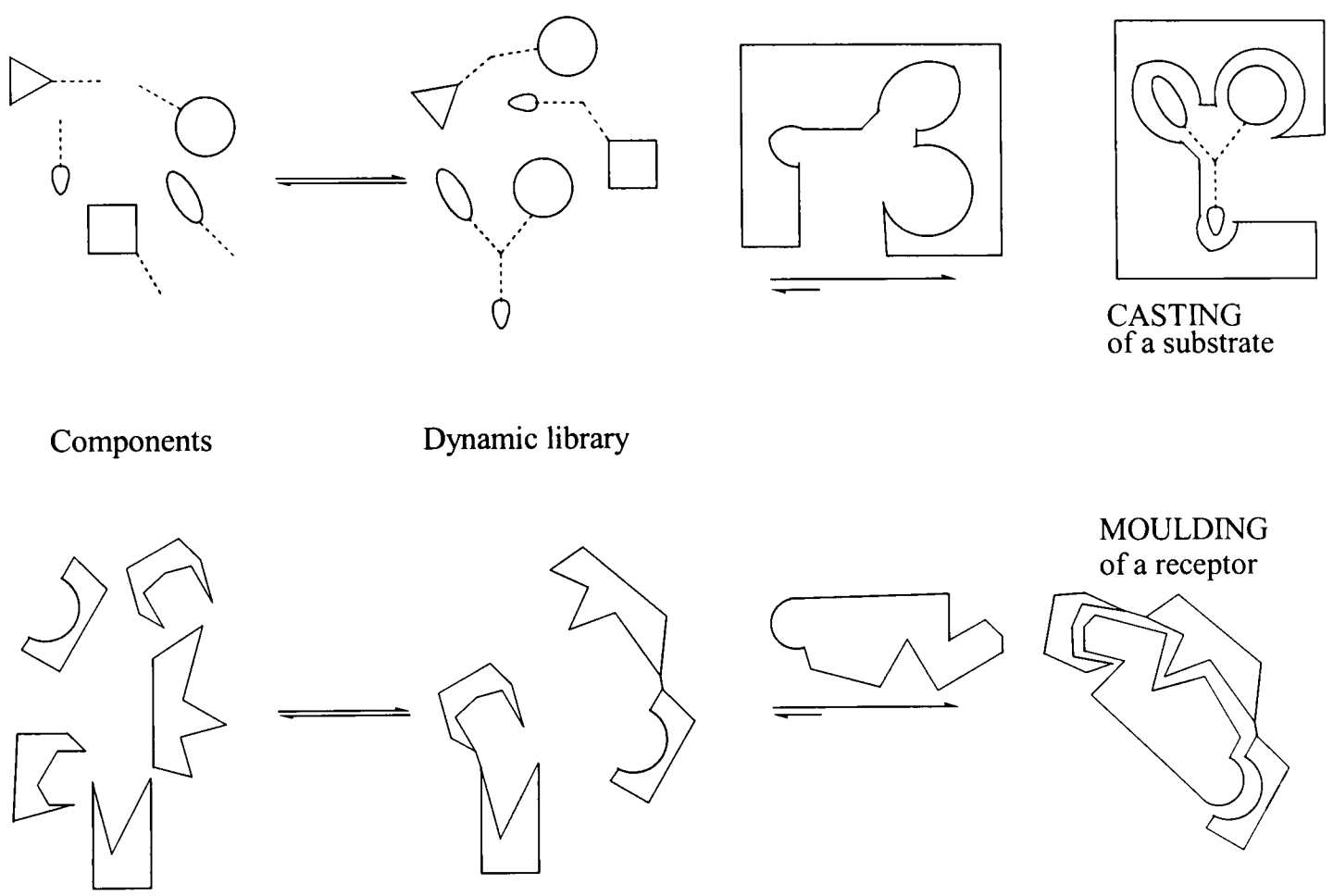
Neither of the native enzymes, nor those that had solely been imprinted with **114**, nor those imprinted and derivatised enhanced the slow hydrolysis of ester **115**. However, when the crosslinked imprinted proteins were used, very impressive relative rate accelerations of 5.4×10^4 mM^{-1} and 1.1×10^5 mM^{-1} were observed for chymotrypsin and subtilisin respectively. This represents the first example where the imprinted property of an enzyme has been stabilised by covalent crosslinking, and therefore been made available for use in aqueous media. Clearly, although still in its infancy, this technique is a very significant advance and holds much promise for the future.

3.2.1.5 Dynamic Combinatorial Libraries.

A dynamic combinatorial library consists of compounds formed by the reversible interactions between a set of building blocks[87]. The proportion of each compound in the library is controlled by thermodynamic equilibria. If a template molecule that binds to one of the library members is introduced, then a selective amplification of the amount of that member could be expected. It therefore follows that molecules identified in a

templated mixture should contain appropriate recognition properties to be able to act as a good host for the template molecule. As with MIP's, if the template is a TSA for a given reaction, then the chances that the host will catalyse the reaction in question should be greater than for an independently designed molecule.

Templating of dynamic combinatorial libraries (DCLs) can be in two forms: casting or moulding (Scheme 26)[87]. In the moulding process, macrocyclic, or macromolecular structures assemble around a template molecule, whilst in the casting process the cavity of an enzyme, protein or other macromolecule induces the assembly of a substrate that fits the cavity. As the system is dynamic, the strong binding of a good receptor/substrate will not only lead to an increase its concentration, but also a decrease in the concentration of the other library members.



Scheme 26

Clearly the advantage of this method of library formation is that of all the possible permutations of products, only the formation of those with the desired binding properties is thermodynamically favoured. There are some disadvantages however. One of the key issues in designing DCLs is the choice of reversible chemical reaction. Thus

far a number of reactions have been tested, including imine exchange[87], isomerisation[88], hydrazone formation[89] and ligand exchange in coordination complexes[90], but few examples of thermodynamic control have been reported. In addition, the termination of the reversible reaction is usually achieved by a change in temperature[91], pH[92], or by chemical methods[87], any one of which can change the structure of the formed ligand or receptor.

This approach has not yet been used successfully in the synthesis of artificial enzymes, but since the concept was only reported in 1996 studies are still in their infancy. This technique has however been used to design enzyme inhibitors, and macromolecules capable of binding templates, and a discussion of these applications makes it easy to see how this approach could be applied in the future to the design and synthesis of artificial enzymes.

By way of illustration, Lehn and Huc reported the use of DCLs to design ligands for the enzyme carbonic anhydrase II[87]. They utilised the imine-aldehyde exchange reaction, and chose the amine and aldehyde library precursors based on the knowledge that *para*-sulfonamides such as **116** are known inhibitors of the enzymes (Figure 18). After equilibration in the presence of carbonic anhydrase II, the library was chemically frozen by reduction with $NaBH_3CN$. Analysis of the library showed that the proportion of *para*-sulfonamide **117** was amplified relative to the library formed in the absence of the receptor.

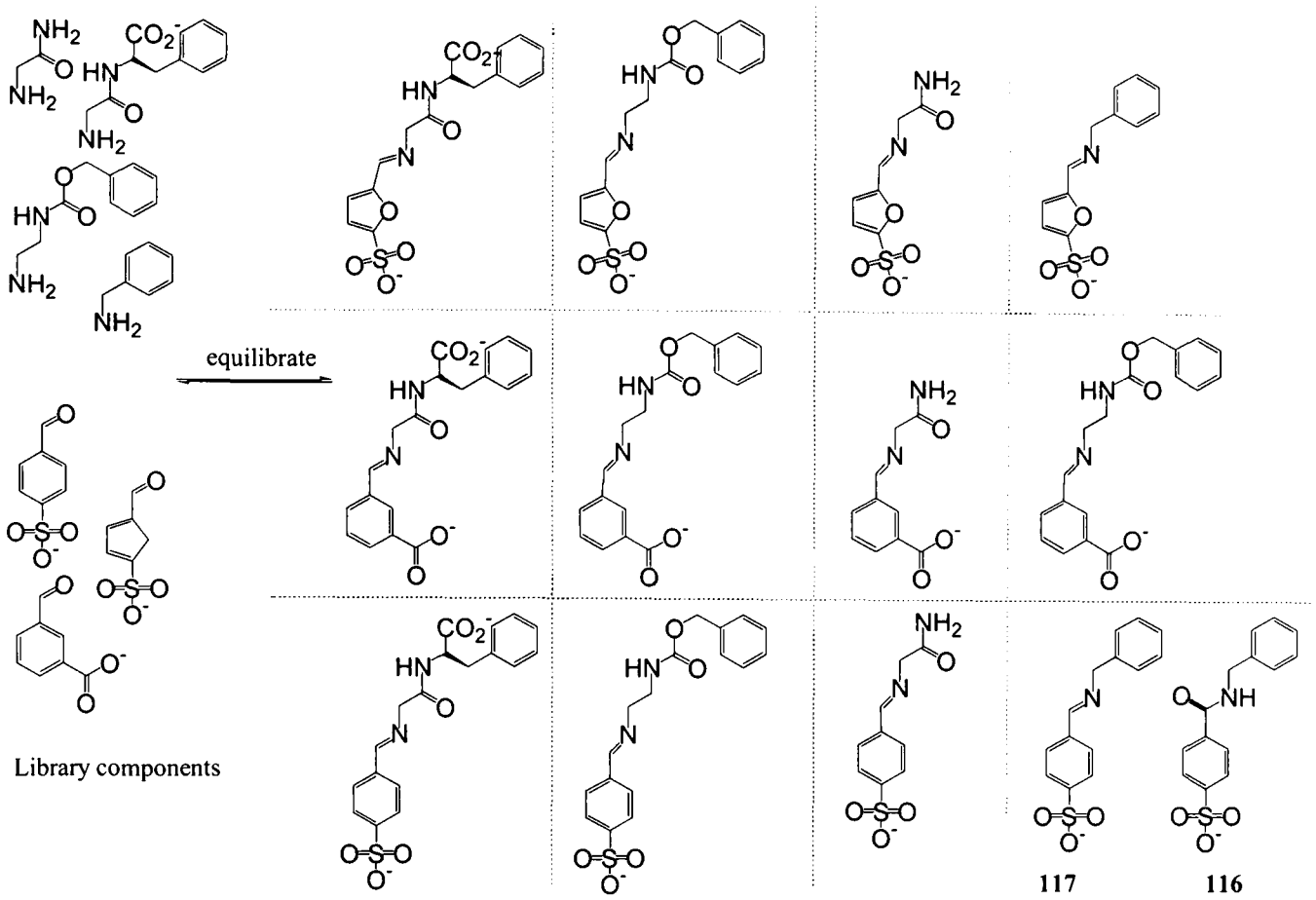
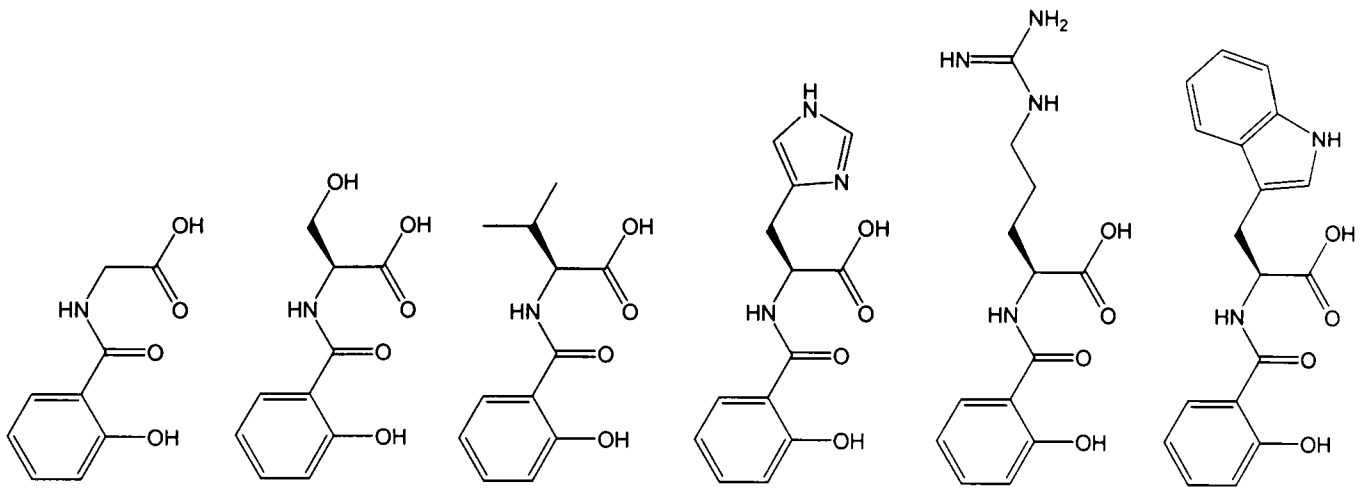


Figure 18

Miller and Karan have recently reported the use of DCLs to identify complexes that selectively bind to RNA[93]. This is a particularly important biological issue, since small molecules capable of binding selectively to RNA over DNA have potential as new therapeutic agents.

Salicylamides **118-123** (Figure 19) were chosen as the components of the DCL, since they would be capable of binding to metal ions. The variable functionality consisted of different amino acid groups as it was predicted these should bind to RNA, since regulation of RNAs by proteins is well documented. Cu(II) was chosen as the metal, since it shouldn't react with the RNA at a typical DCL concentration, and it should be capable of forming spectroscopically detectable structurally well-defined complexes with compounds **118-123**. A library of at least 27 distinct metal complexes should be able to form, assuming the formation of square-planar mono- and bis(salicylamide) complexes.



118 **119** **120** **121** **122** **123**

RNA hairpin **124**

DNA hairpin **125**

Figure 19

The salicylamides **118-123** were equilibrated in a buffer solution, both in the presence and absence of Cu(II) ions. These solutions were then equilibrated either in the presence of the RNA hairpin **124**, the DNA hairpin **125**, or in the absence of any template. As expected it was observed that none of the compounds **118-123** bound to either the DNA or RNA hairpin in the absence of Cu(II). However, analysis of solutions containing Cu(II) equilibrated in the presence of the RNA hairpin **124** showed a significant amount of the complex containing **121** and Cu(II) bound to the RNA. Interestingly, there was a significant difference in the amount of **121** selected by the RNA sequence and the DNA sequence. Further binding studies showed that **121** binds the RNA hairpin with greater than 300-fold selectivity over the homologous DNA sequence.

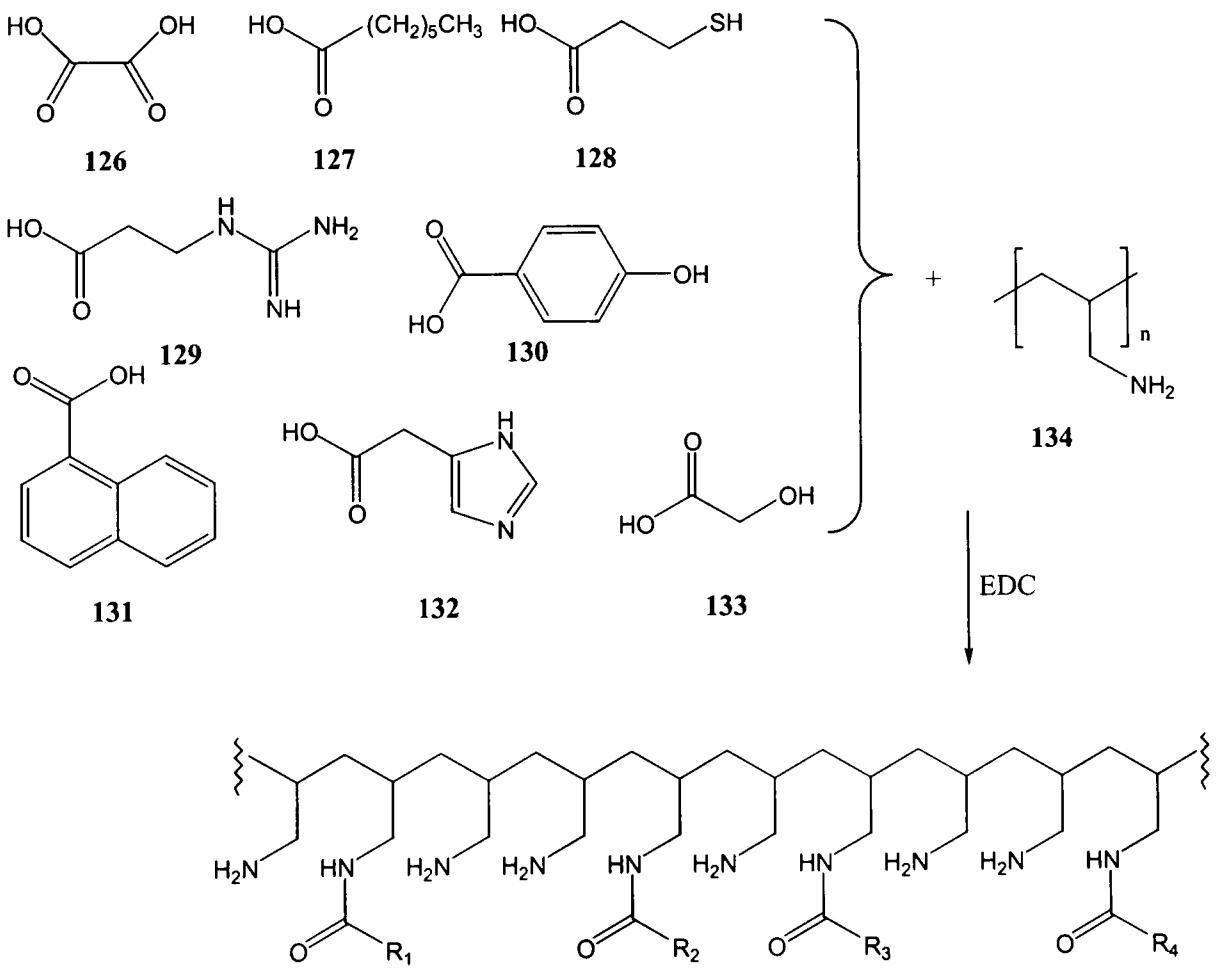
Other recent examples have used DCLs to design receptors for transition metal ions, and have combined the use of different reversible reactions to form double-level DCLs[94]. However, whilst this technique has the potential to be used in the synthesis of artificial enzymes, the first application of this type of approach has yet to be reported.

3.2.2 Catalytic Activity Based Selection Approaches.

The other selection approach uses catalytic activity as the selection event, and utilises combinatorial chemistry and molecular biology to simultaneously synthesise vast libraries of different molecules. The use of combinatorial chemistry in the synthesis of artificial enzymes creates diversity which can be exploited in the search for catalysis by an appropriate choice of screening methods.

3.2.2.1 Combinatorially Generated Polymers As Enzyme Mimics.

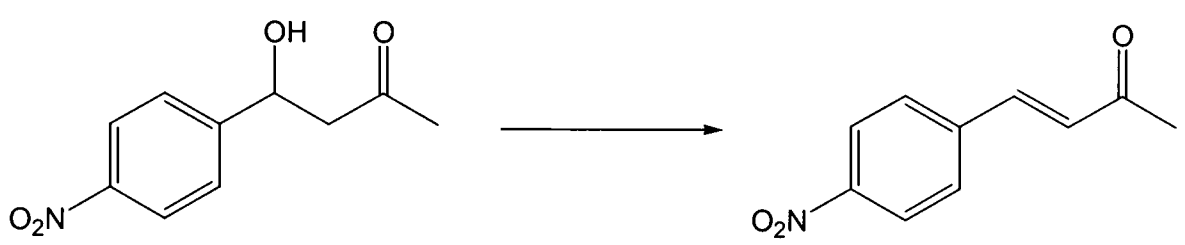
In 1995 Menger used combinatorial chemistry to generate polymer-bound polyamides that acted as catalysts for the hydrolysis of a phosphodiester[95]. A range of different acids **126-133**, in varying amounts were condensed with polyallylamine **134** (Scheme 27). The solutions of the combinatorially generated polymers were then screened for catalytic activity in the hydrolysis reaction with and without Zn(II) ions present. The most active polymer was found to increase the rate of hydrolysis by a factor of 3×10^4. This is over five times more than the best rate enhancement achieved by a catalytic antibody for the same reaction.



Scheme 27

The composition of each polymer is combinatorial, and each individual polymer contains a range of metal ion sites. Thus as well as each polymer differing from each other in composition, each polymer in itself represents a combinatorial range of sites. It is to be expected that there will be regions of local organisation, since one residue attached to the backbone will influence subsequent neighbouring substitutions. Thus, the major drawback of these polymers is that we will never be able to fully characterise the active site. However, these polymers have successfully been utilised not only in the aforementioned reaction, but also in the reduction of benzoylformic acid to mandelic acid[96].

More recently Menger has formed derivatives of poly(acrylic anhydride) using functional amines[97]. The combinatorial polymer formed was capable of catalysing the dehydration of β-hydroxy ketone **135** to α,β-unsaturated ketone **136** with a k_{cat} of 920 over the background reaction (Scheme 28).



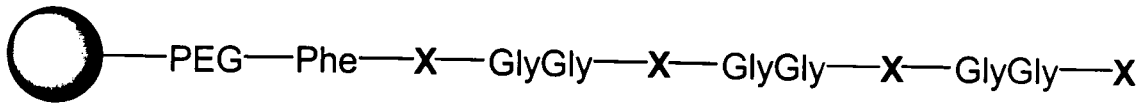
135 **136**

Scheme 28

3.2.2.2 Combinatorially Generated Peptide Catalysts.

Combinatorial chemistry has been used to generate polypeptides for many years. Berkessel recently reported use of a combinatorially developed undecapeptide-zirconium complex as a catalyst for phosphate hydrolysis[98]. A library of 625 solid-phase bound undecapeptides was generated based on the general structure shown in Figure 20. The polymer bound peptides produced contained four variable positions denoted by the letter X, separated by three spacers Gly-Gly and occupied by either L-Arg, L-His, L-Tyr, L-Trp or L-Ser in a combinatorial fashion. Arginine was chosen since in enzymes such as Staphylococcal nuclease which catalyses phosphoryl transfer, the activation of the anioinic phosphate towards nucleophilic attack is by ion-pair formation with the guanidinium moiety of arginine. Both histidine and tyrosine are known to act as ligands for transition metal ions in many enzymes. The aromatic indole system of tryptophan was introduced to display π-π stacking interactions with aromatic substrates, and the

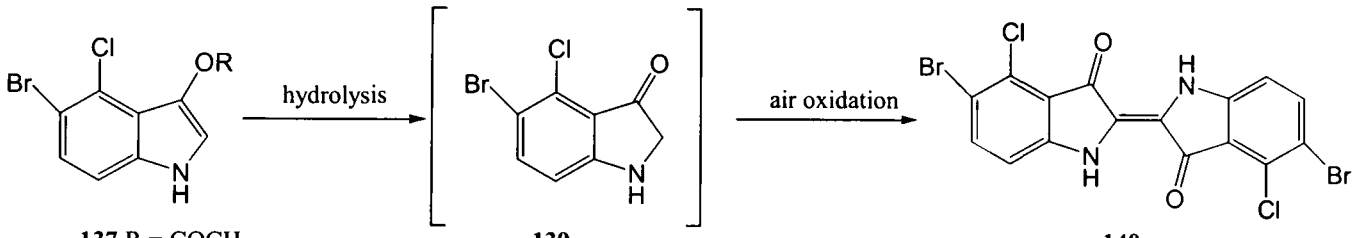
hydroxyl of serine had the potential to serve as the acceptor for a phosphoryl group in a transesterification reaction.



PEG = polyethylene glycol.

Figure 20

The polymer beads were complexed with 1mM solutions of Cu(II), Zn(II), Fe(III), Co(III), Eu(III), Ce(IV) or Zr(IV). The library was screened for hydrolase activity against the 2-hydroxyindole derivatives **137** and **138**. When the ester function is hydrolysed, the resulting indoxyl derivative **139** is oxidised to afford the turquoise and insoluble indigo dye **140** (Scheme 29). Thus any bead which acts as a catalyst will be dyed indigo during the course of the reaction. Activity was found with 20 of the beads in the presence of Zr(IV). Quantitative rate analysis was not possible due to the insolubility of **140**. The three most active, and two of the least active peptides were isolated, and characterised in order to identify any particular residues essential for catalysis. All three peptides that catalysed the reaction contained a serine residue at position X1, and a histidine at X4, suggesting that histidine is essential to bind the Zr(IV) ion, and the serine acts as a phosphate acceptor.



137 R = $COCH_3$
138 R=PO_3^{2-}

139

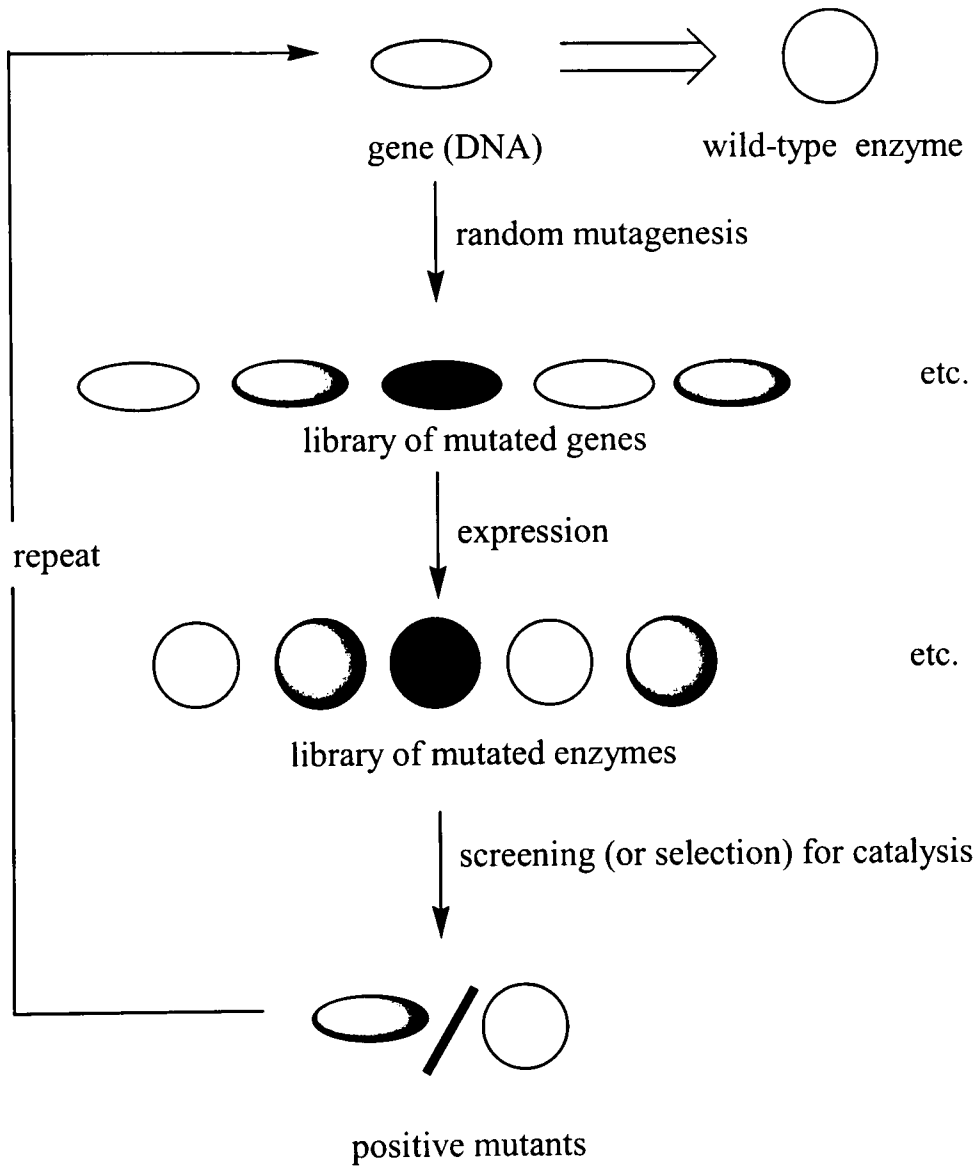
140

Scheme 29

Whilst the employment of peptides as artificial enzymes seems logical, since they contain the catalytic residues found in enzyme active sites, and have the ability to form 3-dimensional helical structures capable of mimicking those found in enzymes, the future of this approach seems limited. Only small peptides can realistically be employed here, since the number of combinations of amino acids in larger peptides is prohibitively large. Combinatorial chemistry has been utilised widely in the search for catalysts, but these are not discussed here, since not many of them come under the umbrella of "artificial enzymes"[99].

3.2.2.3 Directed Evolution of Enzymes.

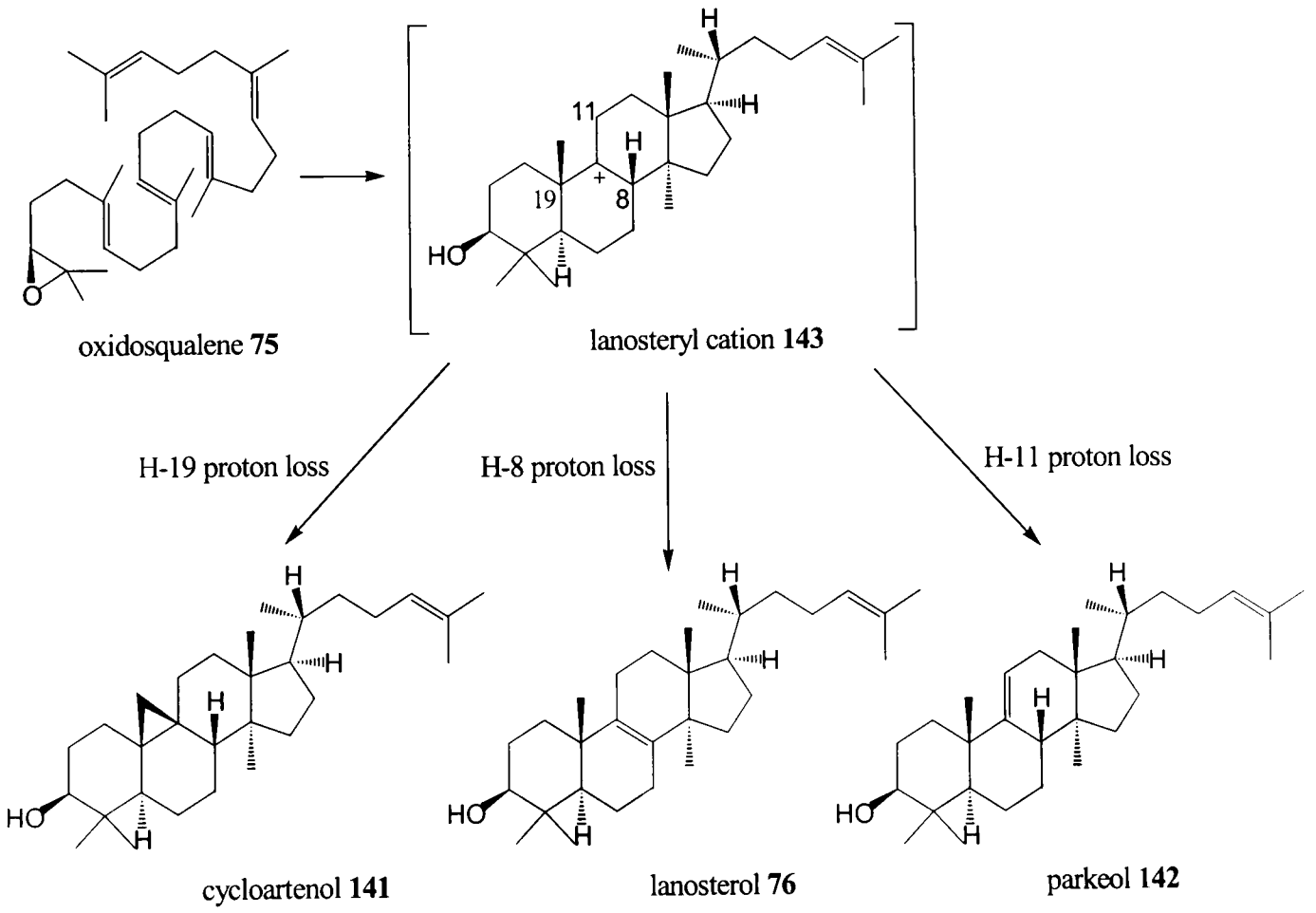
Biochemists have investigated the directed evolution of enzymes in order to find biocatalysts with new or modified activity. Directed evolution strategies generally utilise some method of introducing random mutations into a gene followed by screening or selection for a given property, such as catalysis[101]. The cycle is repeated several times until the selection criteria is achieved, or further cycling produces no further optimisation in the desired property (Scheme 30). Successes have been reported using point mutations generated by a number of techniques such as error prone polymerase chain reaction (PCR) or in more recent examples, DNA shuffling[102].



Scheme 30

The enzyme cycloartenol synthase converts oxidosqualene **75** to cycloartenol **141** and parkeol **142** in a 99:1 ratio *via* an intermediate lanosteryl cation **143**. Hart *et al.* reasoned that mutant enzymes might be obtained that could abstract the C8 proton of the lanosteryl cation **143** to produce lanosterol **76** (Scheme 31)[103]. Lanosterol is a cyclic

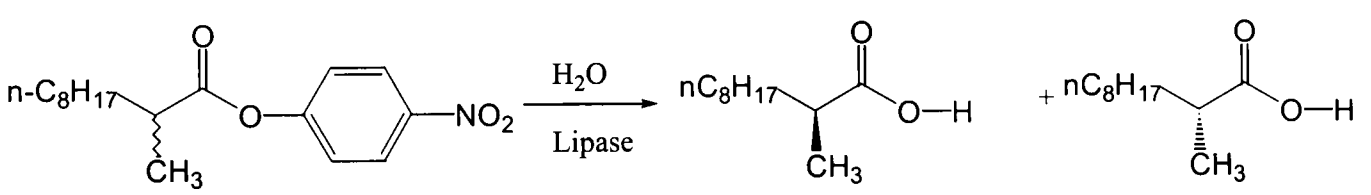
intermediate in the biosynthesis of ergosterol. *Arabidopsis thaliana* cycloartenol synthase was subcloned and expressed into SMY8, a yeast lanosterol synthase deletion mutant that requires ergosterol supplementation. The resultant strain biosynthesises oxidosqualene **75**, which it converts to cycloartenol **141** using the foreign enzyme. The mutant is unable to convert cycloartenol **141** to ergosterol, and therefore remains dependant on exogeneous sterol. The strain was shaken at 30°C in an inducing medium, supplemented with trace amounts of ergosterol. The low ergosterol level limits growth so that after 24 hours the culture was only slightly turbid. After 24 hours, a small aliquot of this culture was used to inoculate more of the same medium, and this process was iterated daily so that rapidly growing mutants that evolved lanosterol biosynthetic ability were able to dominate the culture. After 17 days, the culture reached a high density, indicating the presence of a spontaneous mutant with low ergosterol requirements. When this mutant (Ile481Val) was isolated and used in the cyclisation of oxidosqualene **75**, as well as observing the usual products **141** and **142**, lanesterol **76** was also produced in a 26% yield.



Scheme 31

Aspartate aminotransferase (AspAT) is a homodimeric enzyme that catalyzes amino group transfer between acidic amino acids, aspartate and glutamate, and their

corresponding 2-oxoacids. Oue *et al.* utilised DNA shuffling to change the substrate specificity of AspAT so that it showed activity for β-branched amino acids such as valine[104]. The selection system involved an auxotrophic strain of *E.coli* that is deficient in the gene for branched-chain amino acid aminotransferase. The higher the activity of a plasmid-encoded mutant AspAT for 2-oxavaline, the faster the auxotrophic *E.coli* carrying the plasmid grows on the selection plate. Using this technique Oue isolated a mutant enzyme with 17 amino acid substitutions that showed a 2.1 x 10^6 fold increase in the catalytic efficiency for valine. Interestingly, X-ray crystallography showed that only one of the seventeen mutations appears to be able to contact the substrate, and the role of most of the mutations is in altering the microenvironment, and tertiary structure of the active site. This approach has also been successfully applied by Reetz to synthesise an artificial enzyme which hydrolysed the racemic ester **144** to the *(S)*-isomer with an excellent ee of 81%[105] (Scheme 32).



144 *(S)*-Isomer

Scheme 32

4. Conclusions.

As we have shown above, the techniques used in the design and synthesis of artificial enzymes are continually evolving. The more traditional approach has been the rational design of a molecule utilising our knowledge of enzyme mechanisms. The design tends to be focused on one aspect of enzyme mechanisms, such as designing co-operative catalytic groups, providing a suitable microenvironment for the reaction to occur, and selective binding of a transition state. However, while there have been some considerable successes using rational design, it is recognised that the time and effort involved in affording just one possible host is considerable, and that the smallest flaw in the design can lead to the greatest change in possible activity. Effort is now therefore biased towards selection approaches, where many hosts can be generated at one time. These design studies are worth persevering with however, since even the failures are able to teach us more about the requirements for molecular recognition, and this information is invaluable in the further design of enzyme like systems.

There are two options for selection in the combinatorial approaches to artificial enzymes: binding and catalysis. Selection using binding has been utilised to a greater extent, largely due to it being easier to monitor binding using a TSA than catalysis. However, TSA design is rarely entirely accurate, since the transition state of reaction is a transient species which cannot be isolated. That in designing a stable analogue we can mimic all of the geometric and electronic features of this species is unlikely. It has also been recognised that binding of a transition state is not the only requirement for a molecule to act as an artificial enzyme. Many other factors are involved in enzyme catalysis and, in recognition of this, groups have started to incorporate catalytically active groups into the active site of artificial enzymes. In the introduction of methods such as reactive immunisation in catalytic antibodies, research has started to focus on the design of the library using TSA binding, but selection of hosts based on catalytic activity - a combination which thus far has achieved significant results.

Since the ultimate goal is to design a molecule with catalytic capabilities, then clearly selection by catalytic activity should provide us with the most efficient artificial enzymes. However screening for catalytic activity is not trivial, and it is only due to the

development of new techniques using UV/Vis spectroscopy, fluorescence spectroscopy, IR thermography, and improvements in automated mass spectrometry that this avenue has been opened for exploration. These techniques were initially developed for used in the directed evolution of enzymes, but are now being applied to synthetic combinatorial chemistry. Whilst most success observed in screening for catalytic activity has been in biochemical techniques, synthetic applications have started to appear, as demonstrated with the work by Menger using combinatorial polymers.

Problems that are associated with all selection approaches are the huge number of possible permutations of products possible. This has led to problems in molecular imprinting, when the lack of homogeneity observed in active sites probably leads to a decrease in catalytic activity. In the case of combinatorial polymers, it means that whilst catalysis is observed, we have no way of knowing the actual 3-dimensional structure of the catalyst, which would be essential in trying to improve the efficiency of the catalyst. Combinatorial chemistry is generally performed on very small scales, and isolation and characterisation of the selected hosts can be difficult. In biological techniques this is less problematic, since gene expression leads to amplification of the desired compound. The introduction of DCLs perhaps gives a solution to the synthetic chemist, since the selected host is amplified within the system.

Despite all of these significant advances, there are few examples yet of artificial enzymes that rival the efficiency of natural enzymes. Those that do show significant rate enhancements such as catalytic antibodies, and directed evolution of enzymes still share many of the problems associated with the natural systems, since they too are proteins and share the same lack of thermal stability. Clearly, we still have much to learn from nature itself, and the future of the subject must surely lie in the ability to combine our knowledge with the advances in selection techniques to develop a synthetic receptor that catalyses with efficiency, stereoselectivity and turnover, and can truly be called an "artificial enzyme".

Chapter 2. Results and Discussion.

1. A Novel Semi-Rational Approach to the Design and Synthesis of Artificial Enzymes.

The foregoing introduction has hopefully highlighted that several successful and highly innovative advances have been made in recent years towards the synthesis of effective artificial enzymes. Nevertheless, it is also clear that, whether by the design approach, or through selection, we have still much to learn in terms of formulating an efficient strategy for the creation of these catalysts. All too often, the slightest miscalculation in engineering a tailored 3-dimensional pocket by design can lead to disastrous results, and frustratingly, selection approaches such as those involved in catalytic antibodies[62], or the in vitro evolution protocol pioneered by Reetz[105], whilst leading to active catalysts require retrospective structural analysis to determine the molecule which has been provided.

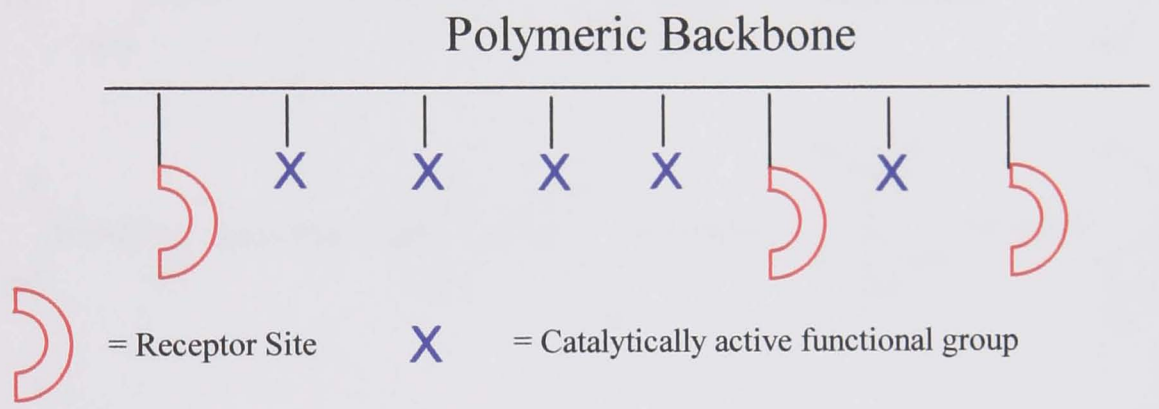
In the light of the above situation, we therefore set out to devise a novel strategy for the modular assembly of artificial enzymes which would not only allow us greater freedom in altering key features such as the receptor site and the catalytically active groups, but also, through systematic investigation, permit us to assess the relative merits of the many and varied factors which are often argued to be essential for enzymatic activity.

Within this framework, we were inspired by consideration of the possible 3-dimensional structure of the inherently flexible polymers produced by Menger in his combinatorial approach (See Chapter 1, p67)[95]. The beautiful and beguiling 3-dimensional structures which are revealed by X-ray crystal structure determination often belie the fact that many enzyme systems can be substantially altered by conformational changes induced on binding. We therefore decided that the greatest chances of success lay in preparing much less rigid systems, and that the capacity for rapid structural alteration through a combination of both design and selection approaches would be the key feature in this novel strategy.

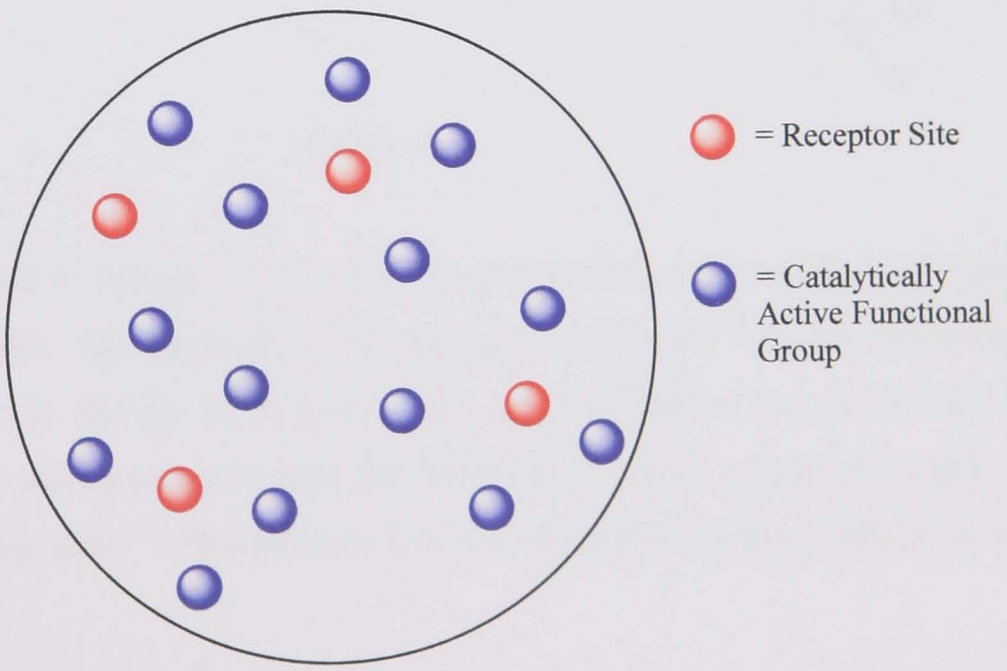
In total contrast to any previous system which has been devised, we decided from the outset that our artificial enzymes would be "physically" divided into a binding "region" for the substrate molecule and a second "region" wherein a catalytically active group

could be located. Whilst the disassembly of a catalytic triad in this manner may be regarded as heresy, our intention was to consider an enzyme at its most simplistic level as a molecule which can both "hold" a substrate (recognition and binding) and also "bite" (carry out a catalytic reaction by virtue of proximal functional groups).

This approach is simplistically shown in Figure 21. In essence, each polymer strand can be regarded as the leg of a millipede which randomly possess either a designed receptor site based on binding a transition state analogue, or a catalytically active group. The exact tertiary structures that evolve will of course be unknown, but the overall flexibility maximises the possibility of developing an effective region for catalysis.



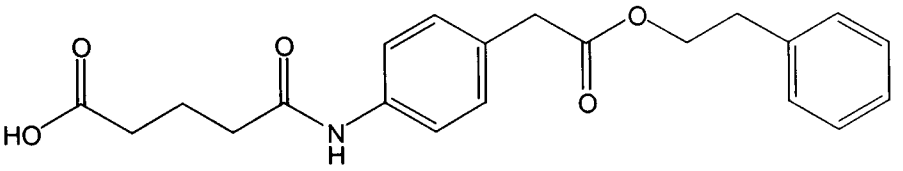
Linear view- Receptor sites are grafted onto a polymeric backbone at random intervals, and possible catalytic groups inserted into sites between the receptor sites



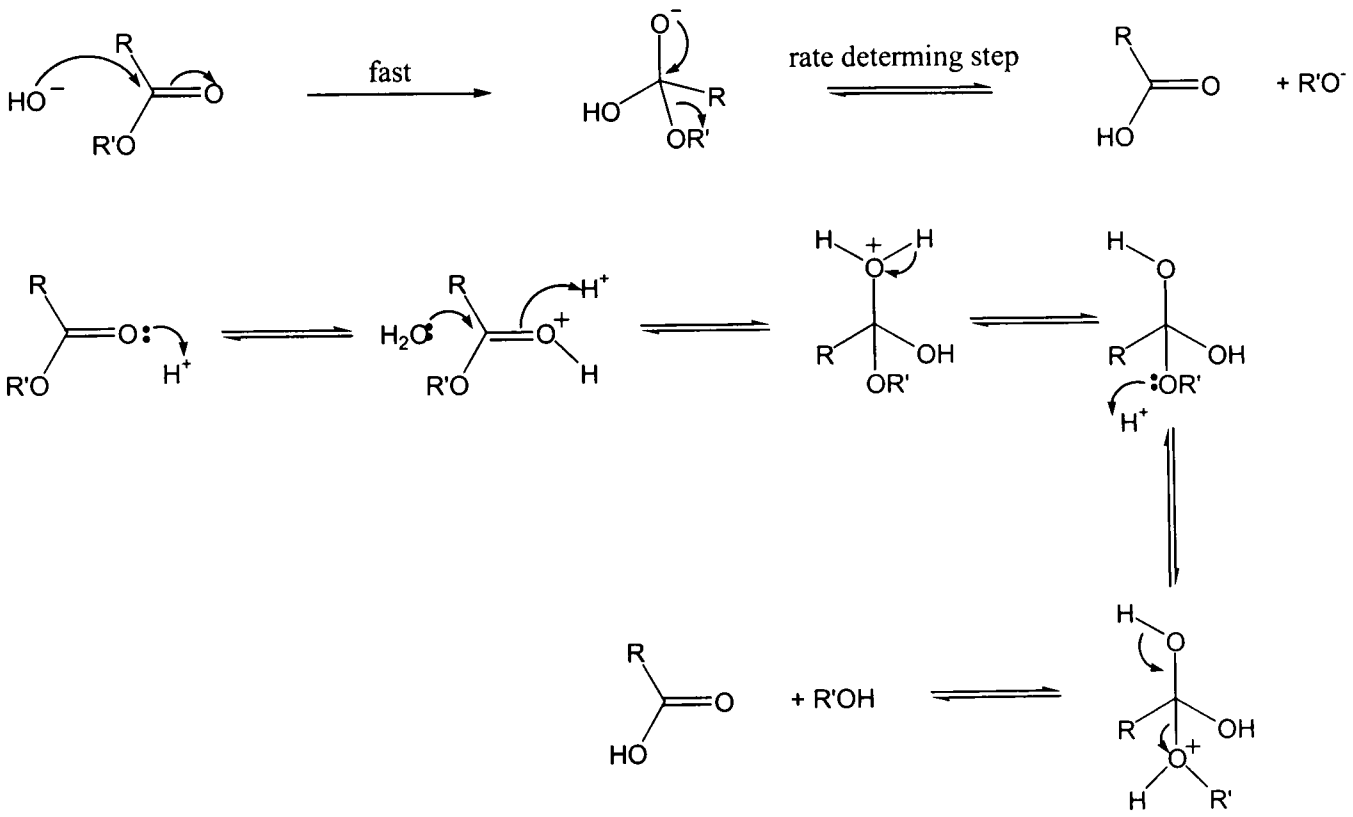
Aerial view- each receptor site is surrounded by at least three catalytic groups, and once a substrate is bound, anyone of these may be in the correct position to attack the substrate and act as a catalyst.

Figure 21

The reaction chosen for study was the hydrolysis of ester **145**. The mechanism of ester hydrolysis is well documented, and proceeds *via* nucleophilic attack on the carbonyl group, resulting in addition, and a change in hybridisation from $sp^2 \rightarrow sp^3$ to give a tetrahedral intermediate which then collapses in the rate determining step to afford the product[106]. The transition state of the reaction is thought to closely resemble the tetrahedral intermediate. Hydrolysis can be catalysed by either acid or base (Scheme 33).

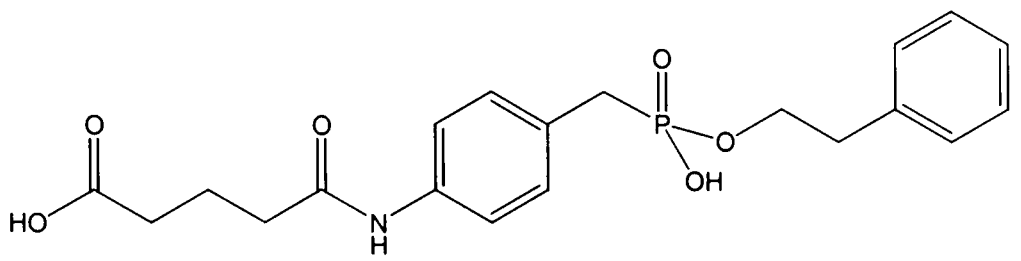


145



Scheme 33

As the transition state is thought to resemble the tetrahedral intermediate, any stable analogue that mimics the geometry and charge distribution of this tetrahedral intermediate should in theory be a good TSA. Phosphonates have been used as inhibitors of natural esterases, since they are believed to exhibit properties similar to those of the transition state[107]. Phosphonate **146** was therefore a logical choice for our TSA (Figure 22).



146

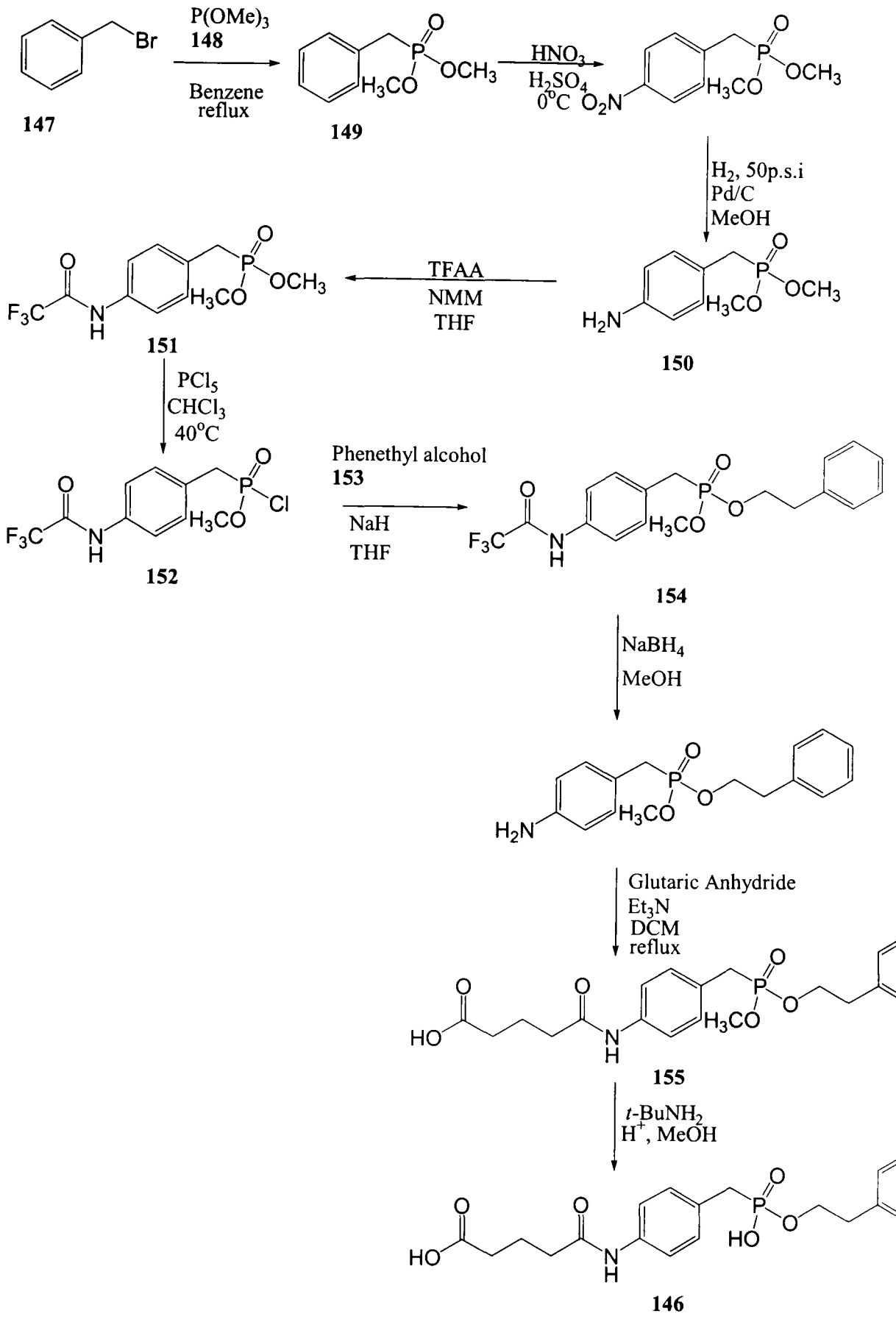
Figure 22

This specific ester was chosen for a number of reasons. Firstly the TSA **146**, and hence the transition state of the reaction contains three possible binding sites. Both the carboxylic acid and phosphonate moieties should be able to bind to bases, and hydroxyl groups, and the amide linkage has the potential to form hydrogen bonds with other polar groups. Secondly, all three of the binding sites are on the acid side of the ester bond, and therefore the catalyst formed may act on a range of esters, as the alcohol could be varied. Finally, an unactivated ester was chosen so that any rate enhancement could be seen as a significant result.

2. Synthesis of the Transition State Analogue.

A nine-step synthesis of the TSA **146** starting from benzyl bromide **147** and trimethyl phosphite **148**, both cheap commercially available products was selected (Scheme 34). Thus, a Michaelis-Arbuzov reaction between **147** and **148** should yield the dimethyl benzyl phosphonate ester **149**. Nitration and reduction then lead to the *p*-amino compound **150**, and subsequent protection using the trifluoroacetyl group should give compound **151**. Monochlorination forms a reactive species **152** that, under basic conditions should react with alcohol **153** to form the mixed phosphonate **154**. Chemoselective deprotection of the amide and reaction with glutaric anhydride then furnishes the amide **155**. Finally selective demethylation of the phosphonate with either tertiary butylamine or trimethylsilyl bromide yields the required TSA **146**.

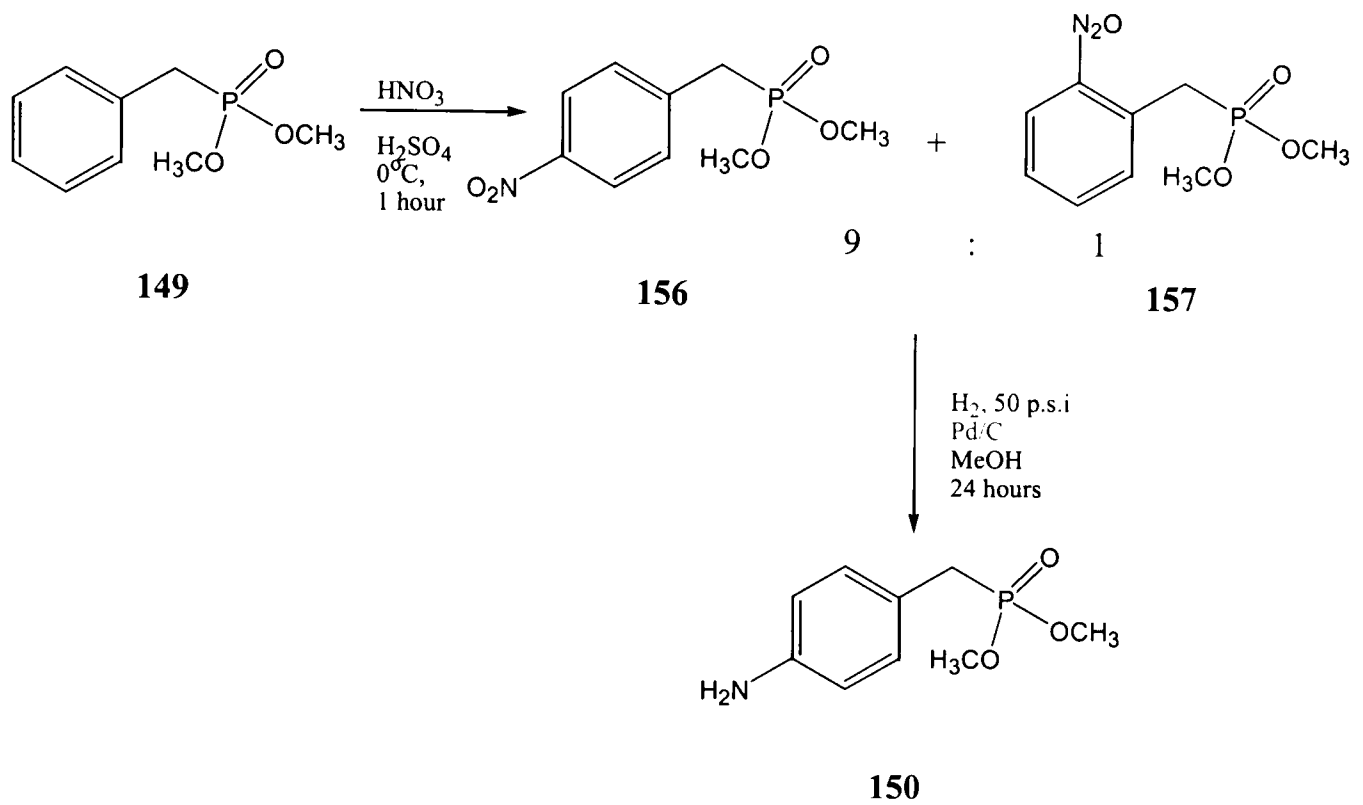
The initial step of the synthesis was the Michaelis-Arbuzov reaction between benzyl bromide **147** and trimethyl phosphite **148**. The *para*-nitrated benzyl bromide could not be used due to possible interference of the nitro group, and attempts by other groups to perform Arbuzov reactions with nitrated compounds have resulted in failure[108]. Following a literature procedure, benzyl bromide **147** was treated with 10 equivalents of trimethyl phosphite **148** in refluxing benzene. Removal of excess reagents, and distillation afforded dimethyl benzylphosphonate **149** in a 93% yield.



Scheme 34

At this stage we wished to achieve selective *para*-nitration of the benzene ring. Although treatment of phosphonate **149** with a nitric acid-sulphuric acid mixture at 0°C yielded a nitrated product in a 70% yield, NMR analysis revealed that the product contained both the *ortho* **156** and *para* **157** isomers in a 1:9 ratio (Scheme 35)

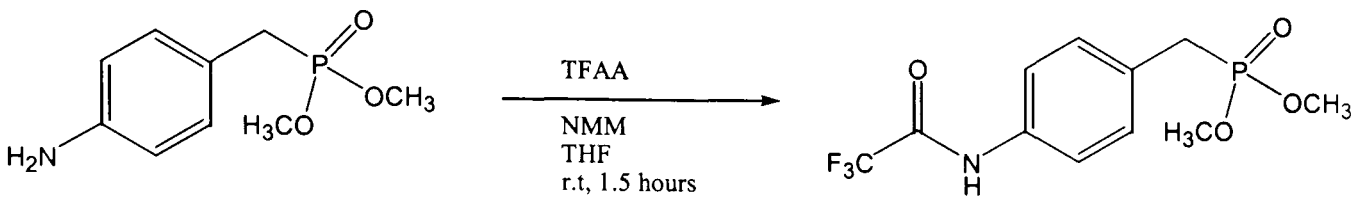
Unfortunately, these could not be separated. It is presumed that the *para* isomer favoured due to the steric bulk of the phosphonate group hindering nitration in the *ortho* position.



Scheme 35

Fortunately catalytic hydrogenation of **156** in the presence of 10% palladium on charcoal led to selective reduction of the *para*- isomer to afford *p*-aminobenzylphosphonate **150** in 57% yield. Initially this figure seemed poor, but in fact compared well with the 55% yield reported in the literature[109].

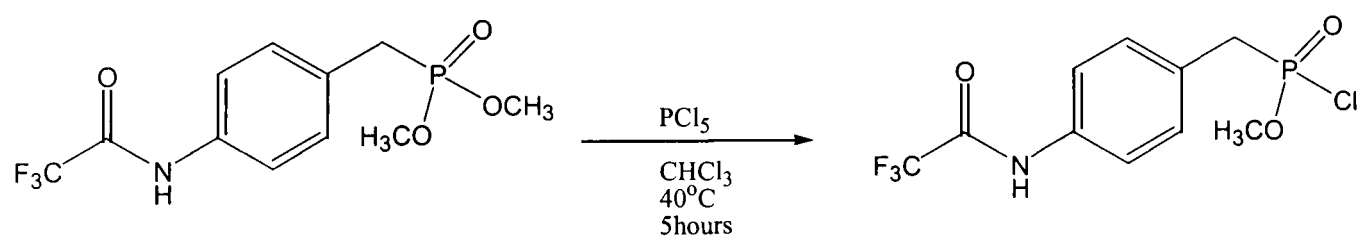
Protection of the amino group was then required to prevent possible reaction of the amine with the phosphonochloridate formed in the next step. The protecting group chosen had to be stable under the acidic conditions required to form the chloride species, and also under the basic conditions required to form the mixed phosphonate ester. The trifluoroacetyl group fulfilled these requirements, with the added bonus of being easily reduced using sodium borohydride[110]. Treatment of the amino compound **150** in anhydrous THF with trifluoroacetic anhydride and *N*-methylmorpholine as described by Moree *et al.* yielded the required protected amine **151** in an 80% yield after chromatography (Scheme 36)[111].



150 **151**

Scheme 36

The dimethyl phosphonate **151** had to be converted into a reactive species for coupling with phenethyl alcohol **153**. The most common reactive intermediate utilised is the phosphonochloridate species, which is thought to behave in an analogous manner to an acyl chloride[112]. The tried and tested method for conversion of a dialkyl phosphonate to the mono phosphonochloridate has been by treatment of the ester with phosphorus pentachloride[112]. Following this procedure, **151** was treated with phosphorus pentachloride in chloroform and heated to 40°C for 5 hours. The product can be observed using ^{31}P NMR as a peak at 41 ppm, with the starting material at 30 ppm. After 5 hours ^{31}P NMR showed no remaining starting material, and good conversion to the required product **152**, however several unknown by-products with peaks between 30 and 35 ppm also seemed to have formed (Scheme 37).

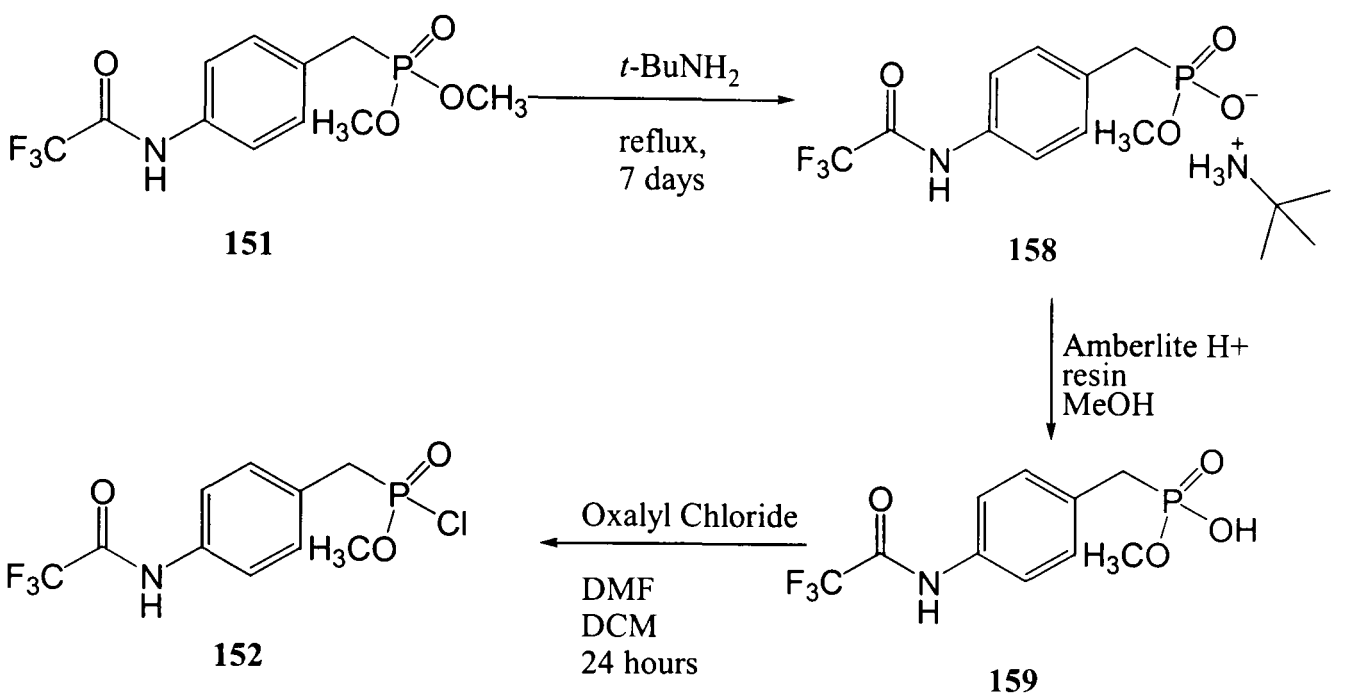


151 **152**

Scheme 37

Methyl phosphonate esters are susceptible to methylation by an S_{N2} mechanism, and it was a concern that the phosphonochloridate formed may be sensitive to any residual PCl_5, and other by-products formed during the course of the reaction. Separation of the product from the phosphorus waste also proved to be problematic. It was therefore decided to try a milder approach as described by Coward[113]. In this approach, the dimethyl phosphonate is first converted to the monomethyl phosphonic acid by treatment with *tert*-butylamine and then treated with oxalyl chloride to form the

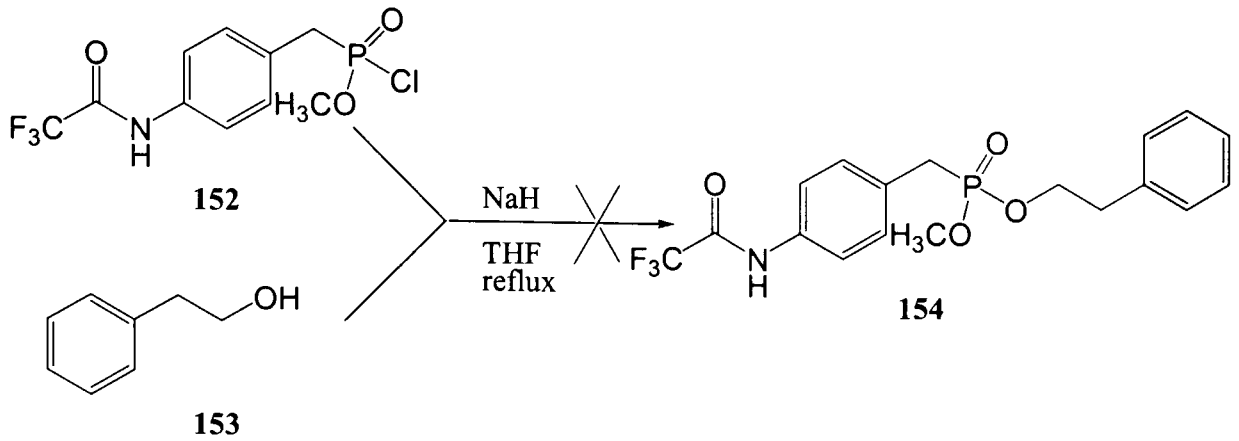
phosphonochloridate species (Scheme 38). The mechanism of the demethylation step has not been discussed in the literature[113,114]. The overall yield of the reaction has been reported to be slightly lower than that obtained with PCl_5, but this method was preferred because of the volatile nature of oxalyl chloride, and thus its easy removal after the reaction had gone to completion.



Scheme 38

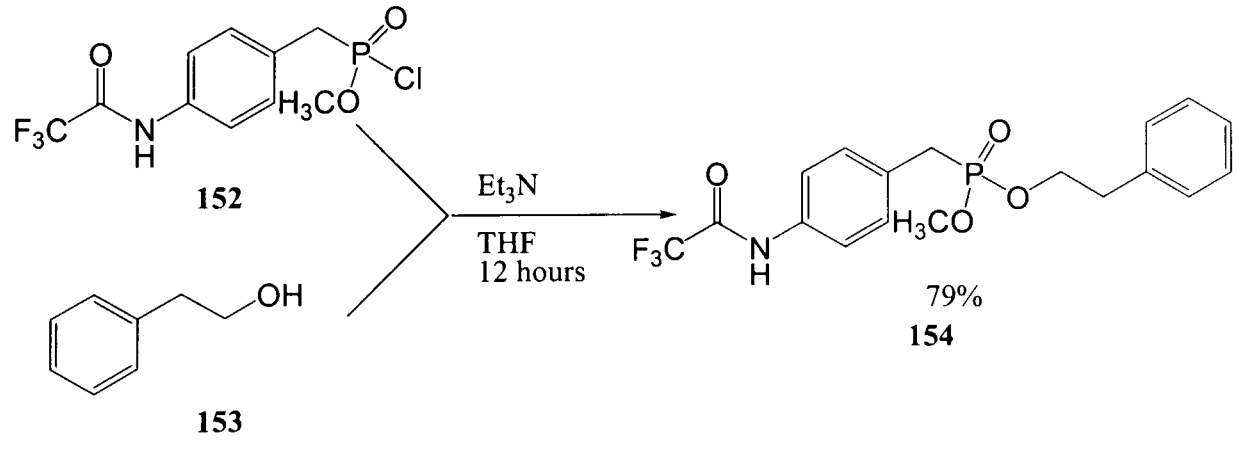
Dimethyl phosphonate **151** was treated with an excess of *tert*-butylamine and heated to reflux. After 7 days TLC showed the reaction to have gone to completion, and removal of the excess amine yielded the *tert*-butylamine salt **158** in a 92% yield. The free phosphonic acid **159** was obtained in 68% yield by stirring with a cation resin in methanol for four hours. Finally the phosphonochloridate **152** was obtained after treatment of **159** with oxalyl chloride and a catalytic amount of DMF in dry DCM. ^{31}P NMR showed clean conversion to the product in 80% yield.

The phosphonochloridate had now to be coupled with phenethyl alcohol **153** to form the mixed phosphonate **154**. In the synthesis of a similar TSA, Lerner had obtained the mixed phosphonate by treatment of a phosphonochloridate with sodium hydride and an alcohol[115]. When the same conditions were tried with our phosphonochloridate however, none of the desired product **154** was formed (Scheme 39).



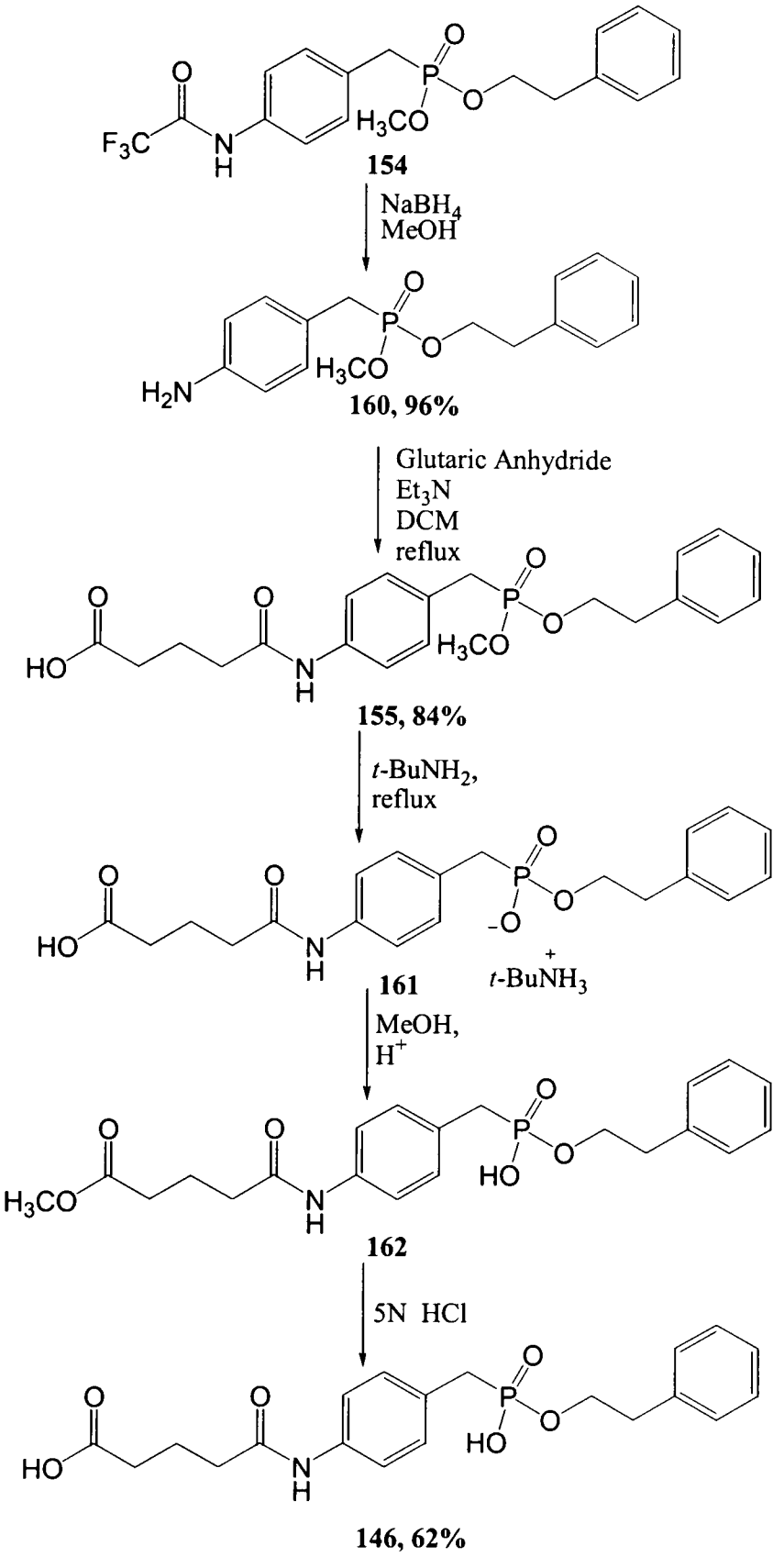
Scheme 39

The use of different bases has been reported in the coupling of phosphonochloridates and alcohols. Coward reported the use of 2 equivalents of triethylamine in THF to successfully form mixed phosphonates[114]. THF is used as the solvent here as it precipitates the triethylamine hydrochloride formed during the course of the reaction, and thus pushes the reaction towards formation of product. When this approach was tried using **152** and phenethyl alcohol **153**, the required product **154** was obtained in a 79% yield after chromatography (Scheme 40).



Scheme 40

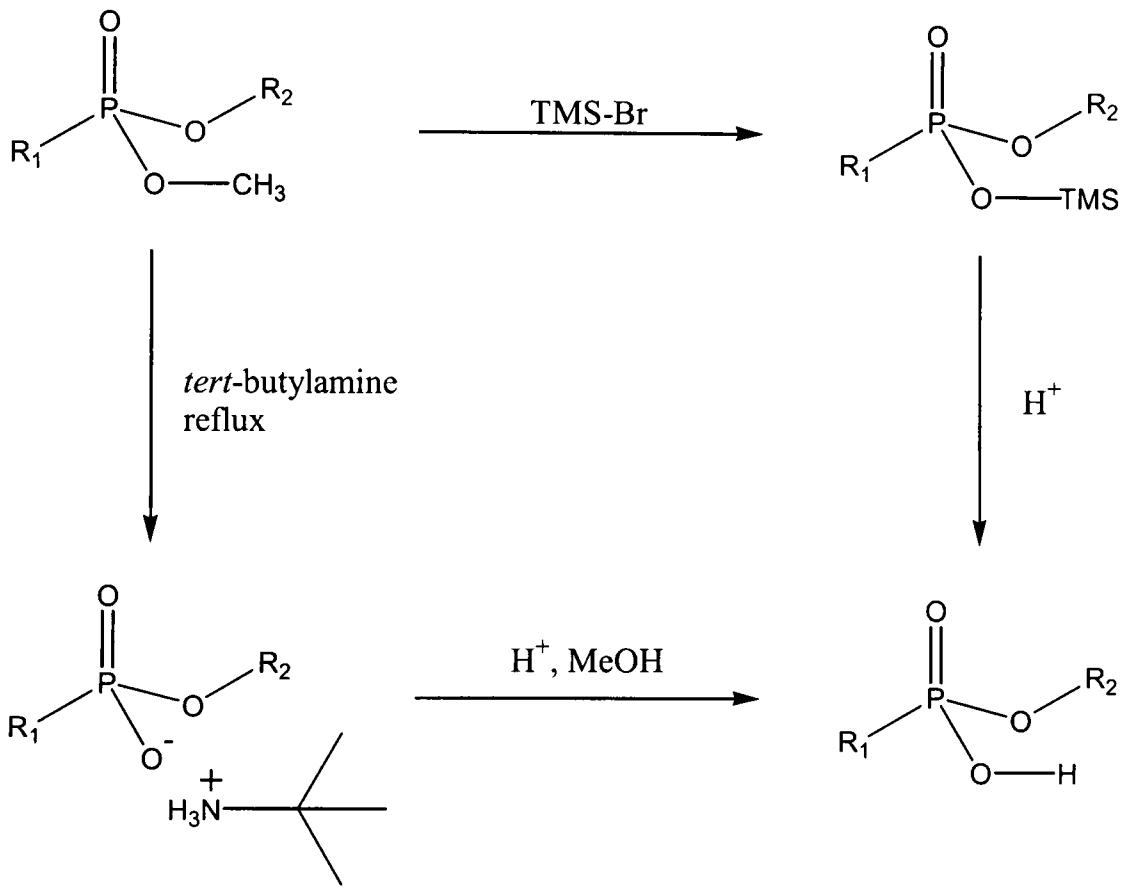
Deprotection of the trifluoroacetyl group can be effected either by treatment with aqueous sodium carbonate solution, or by reduction with sodium borohydride[110]. Since no other functionality within the molecule was likely to be affected by the presence of sodium borohydride the decision was made to use the latter. When **154** was treated with 10 equivalents of sodium borohydride the deprotected amine **160** was obtained in 96% yield (Scheme 41).



Scheme 41

A method described by Lerner was then used to incorporate the amide side chain of **146**[115], and involved the addition of glutaric anhydride to a solution of triethylamine and the *p*-aminophosphonate **160** in dry DCM The solution was then refluxed for 12 hours to afford the product **155** in an 84% yield (Scheme 41).

Two methods were considered for the selective demethylation of the mixed phosphonate **155**. Trimethylsilyl bromide selectively reacts with methyl phosphonate esters to form an intermediate TMS ester, which can be hydrolysed under mild conditions to afford phosphonic acids (Scheme 42)[116]. The other method was the one utilised earlier, in which *tert*-butylamine selectively deprotects methyl phosphonate esters.

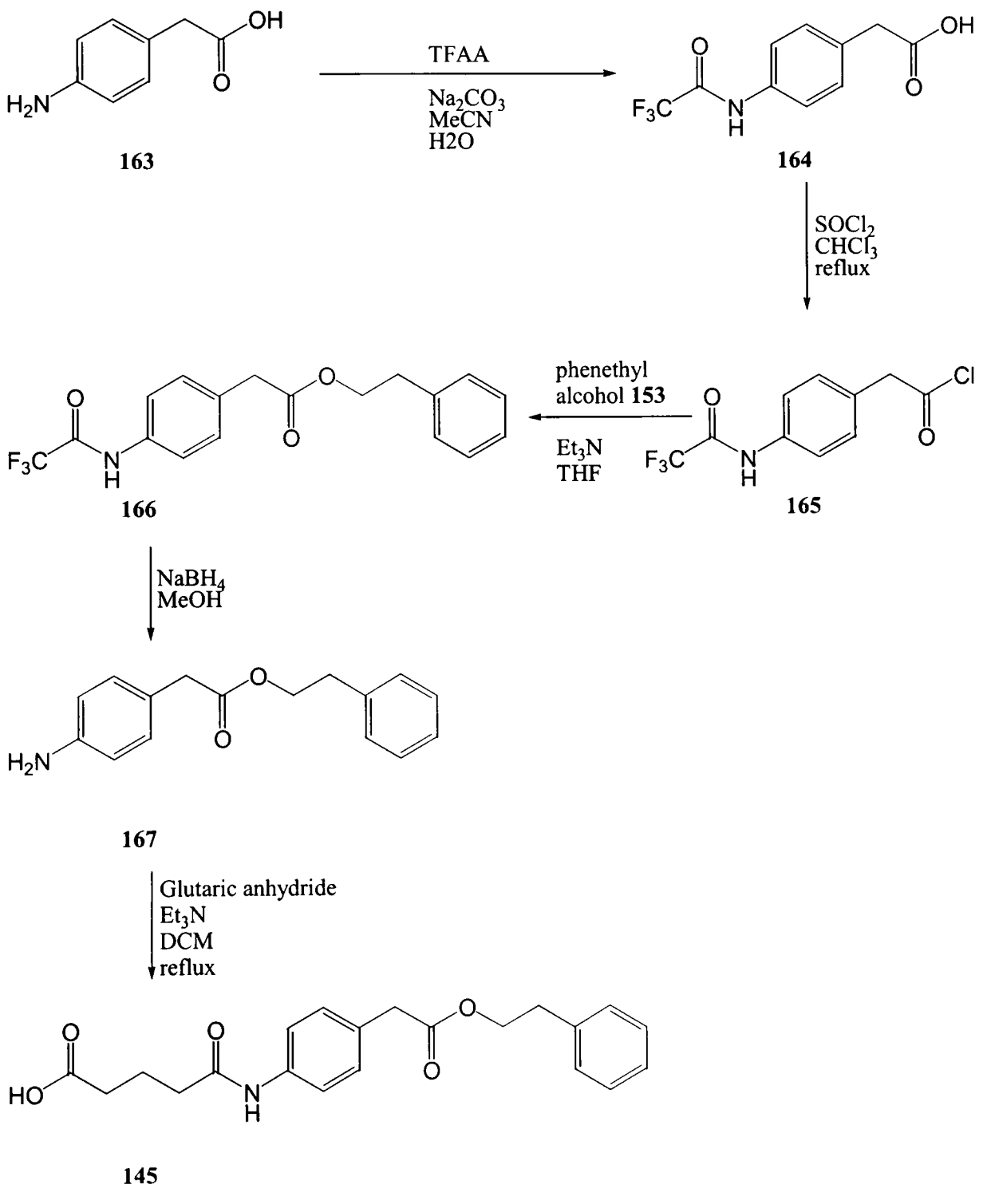


Scheme 42

The *tert*-butylamine method was preferred, since purification of the final product occurs when the product is stirred with the acid exchange resin. The *tert*-butylamine salt **161** was obtained after stirring in the refluxing amine for 7 days. Unfortunately however, stirring **161** in methanol with a cation exchange resin did not yield the required TSA molecule **146**. Instead, ^{1}H NMR showed that formation of the methyl ester **162** had occurred. Hydrolysis of this methyl ester proved facile by treatment with 5N HCl to afford the required TSA **146** in 62% yield over the two steps (Scheme 41).

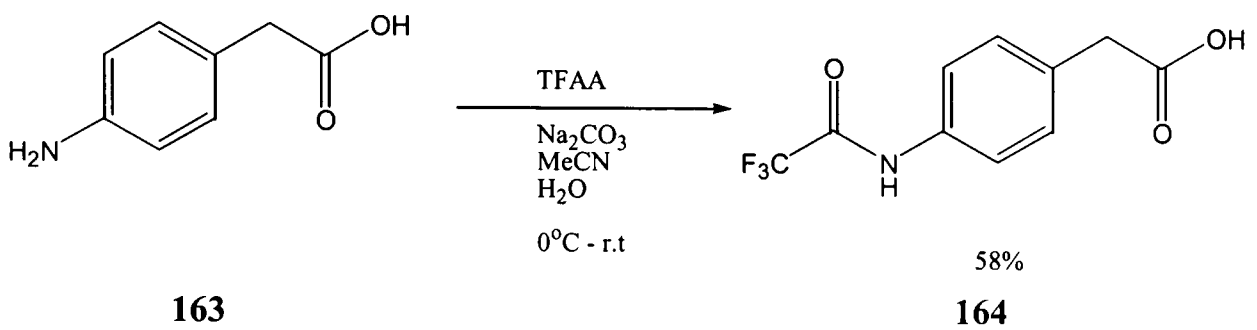
3. Synthesis of the Ester Substrate.

A five-step synthesis of ester **145** was devised starting from commercially available *p*-aminophenylacetic acid **163** (Scheme 43). Initial protection of the amine as its trifluoroacetamide gives **164**. Formation of the acid chloride **165** followed by esterification with phenethyl alcohol should afford ester **166**. Subsequent deprotection of the amine would give **167** which followed by coupling with glutaric anhydride then provides the desired ester substrate **145**.



Scheme 43

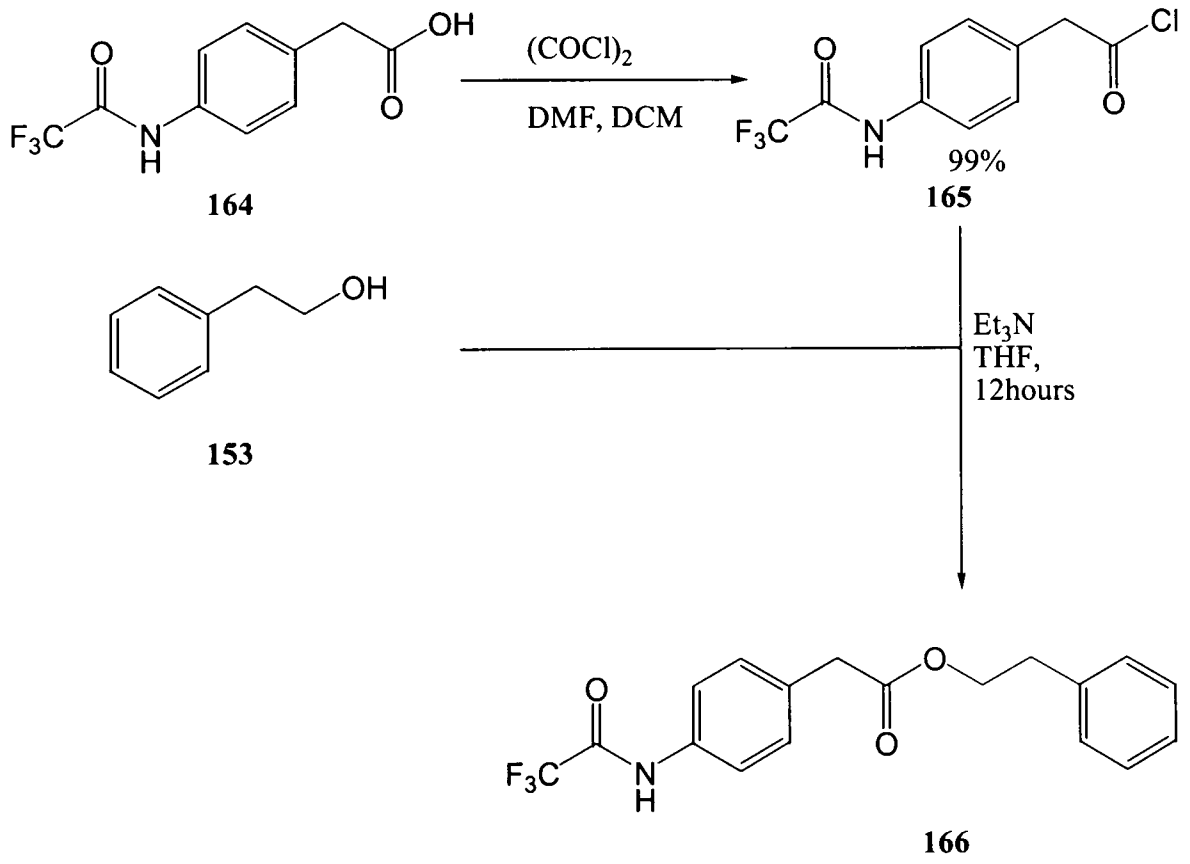
As in the synthesis of the TSA **146** the trifluoroacetyl group was used to protect the amine. Treatment of *p*-aminophenylacetic acid **163** with trifluoroacetic anhydride and sodium carbonate in a mixture of acetonitrile and water afforded the protected amine **164** in 58% yield (Scheme 44).



163 **164**

Scheme 44

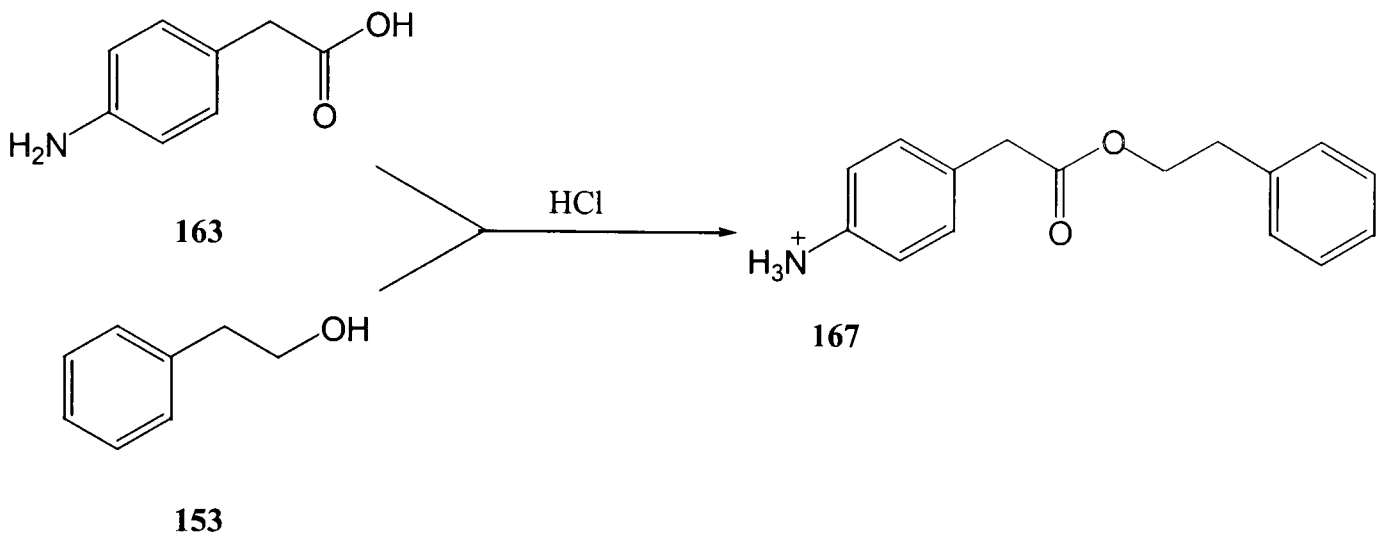
The acid was activated for the esterification step by forming the acid chloride. Protected amine **164** was suspended in anhydrous DCM and stirred with oxalyl chloride, and a catalytic amount of DMF. When all the acid had been converted to the chloride **165**, the solution became homogeneous. The acid chloride **165** was immediately reacted with phenethyl alcohol **153** and triethylamine in anhydrous THF to afford, after chromatography, the required ester **166** in 63% yield (Scheme 45)



164 **165** **153** **166**

Scheme 45

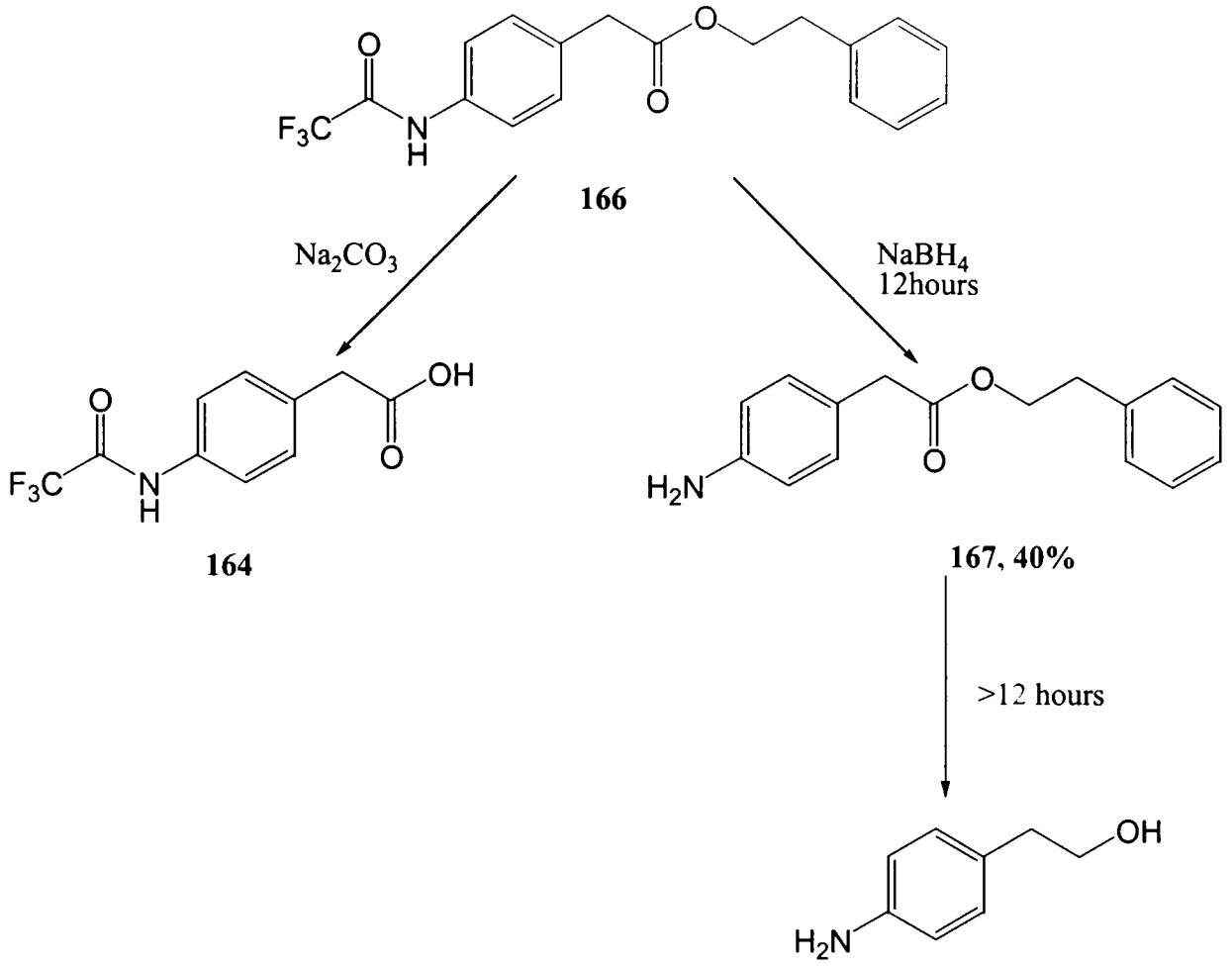
Esterification of amino acids is generally effected by treatment with an excess of alcohol and HCl gas. This approach was attractive as it would by-pass the need for an amine protecting group. Therefore, *p*-aminophenylacetic acid **163** was dissolved in an excess of phenethyl alcohol **153** and treated with HCl gas (Scheme 46). The product **167** was formed, however separation proved troublesome and so this approach was abandoned.



163 **153** **167**

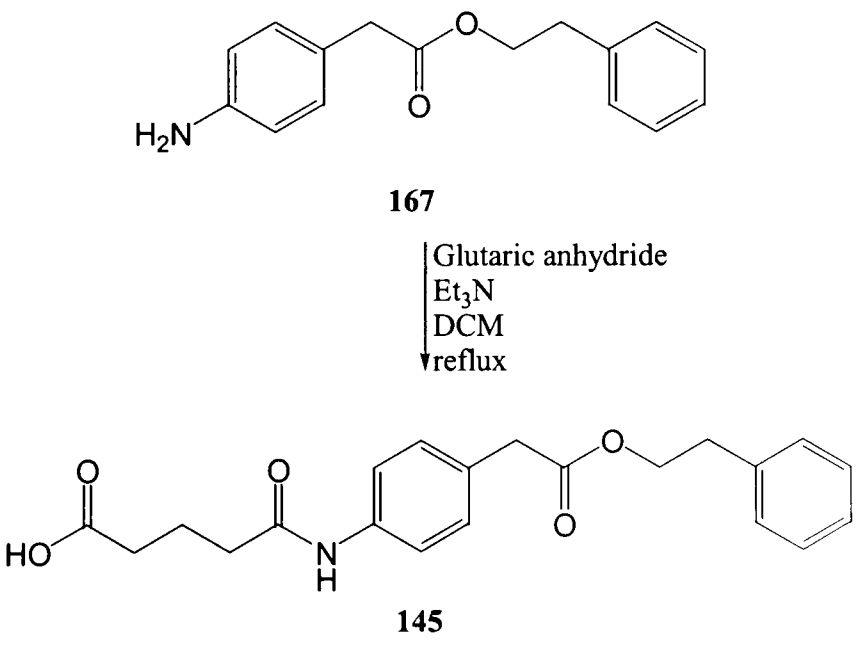
Scheme 46

Deprotection of trifluoroacetyl groups is usually effected either by hydrolysis with sodium carbonate, or by reduction with sodium borohydride[110]. Initially sodium carbonate was utilised, since the activated trifluoroacetamide group should be more prone to hydrolysis than the stable ester. However, when **166** was treated with sodium carbonate, hydrolysis of the ester occurred in preference to the amide group. Sodium borohydride is known to preferentially reduce amides over esters, unless in the presence of a Lewis acid. When **166** was treated with 5 equivalents of sodium borohydride and stirred for up to 12 hours, the required deprotected amine **167** was obtained albeit in a 40% yield. If the reaction was left to stir longer, or if more equivalents of sodium borohydride were added then further reduction of the product **167** started to occur preferentially over the starting amide **166** (Scheme 47).



Scheme 47

The final step in the synthesis was to couple the amine with glutaric anhydride. Gratifyingly, treatment of the amine **167** with glutaric anhydride and triethylamine in DCM afforded the required ester **145** in 99% yield (Scheme 48).

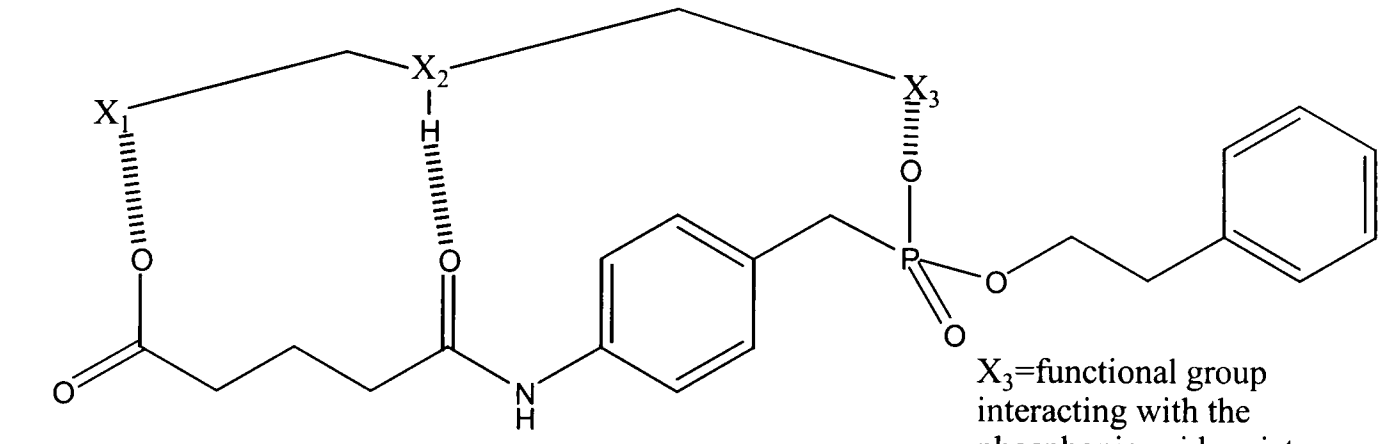


Scheme 48

4. Design of the Receptor Site.

4.0 Possible Receptors.

With the TSA **146** in hand, we were ready to design our "receptor site" which would be incorporated into the polymer with the role of stabilising the developing tetrahedral transition state of the reaction. The receptor site would be designed and selected based on its ability to bind to the TSA. The TSA **146** had been chosen such that there were three possible points for recognition (Figure 23). The receptor site was envisaged as consisting of one molecule with the potential to bind the TSA both at the carboxylate, and the phosphonic acid moieties, and also to form a hydrogen bond with the amide linkage.



146

Figure 23

In theory, the receptor site could have been constituted from a number of discrete entities, with each one binding to one of the aforementioned points on the TSA. However, if more than one functional monomer is used to achieve these multiple binding interactions, the complex becomes less entropically favourable[118]. An additional advantage of the single molecule receptor site was that it would then be possible to study the binding between the receptor and the TSA prior to its incorporation into the

polymer. If three distinct binding entities were used, then they would need to be correctly positioned on the polymer before the binding interactions could be understood.

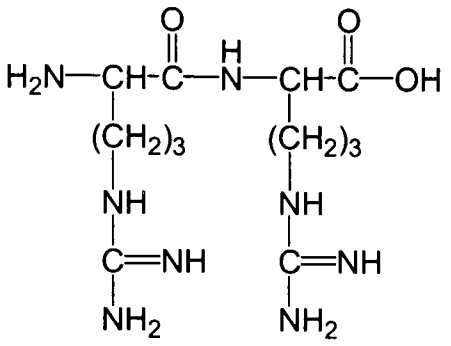
The Cambridge Structural Database[166] was used to highlight different functional groups that show favourable interactions with the different moieties of the TSA. As expected, carboxylic acids form complexes with basic groups. Phosphonic acids also form energetically favourable complexes with bases, but also several complexes have been crystallised in which the phosphonic acid interacts with hydroxyl groups. Finally, as expected, aromatic aliphatic amides, such as the one in our TSA, form hydrogen bonds with other amide groups.

Natural amino acids contain functionality that should bind to the TSA. In addition, we reasoned that if a dipeptide were chosen with suitable functionality to bind to the carboxylic acid and phosphonic acids moieties, the peptide bond should also be appropriately placed to form a hydrogen bond with the amide in our TSA. It was therefore decided to choose the receptor unit from a series of dipeptides (Figure 24).

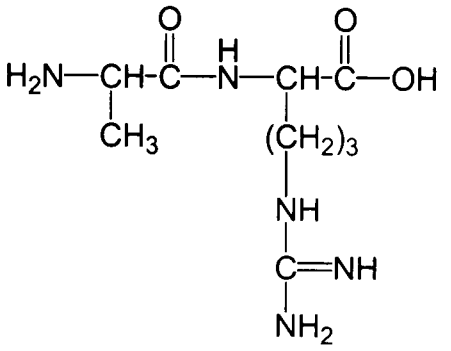
Nine commercially available dipeptides were chosen for use in binding studies with the TSA (Figure 24). Four, Arg-Arg **168**, Ala-Arg **169**, β-Ala-Lys **170** and L-Carnosine (β-Ala-His) **171** contained one or more basic groups. Two dipeptides, Ala-Tyr **172**, and Ala-Thr **173** contained hydroxyl residues, and one dipeptide, Ser-His **174** contained both basic and hydroxyl residues. For comparison in the binding studies, two peptides, Ala-Gly **175** and β-Ala-Leu **176** which contained neither basic, nor hydroxyl groups were included in the set.

A variety of techniques have been reported in the literature for studying binding interactions between a ligand and host. The majority use NMR methods, but mass spectrometry techniques such as MALDI-MS and ESI-MS, as well as affinity chromatography have also been utilised. A recent review highlights the advances and relevant applications in all of these areas[119]. However, NMR methods seemed most appropriate to look at the binding between the TSA and peptides, since all the studies could be performed first hand in house.

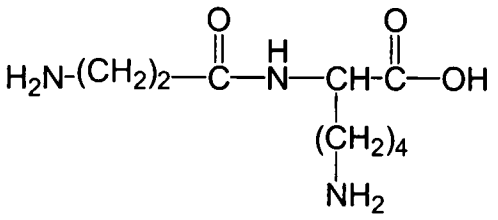
Dipeptides containing 1 or more basic residues



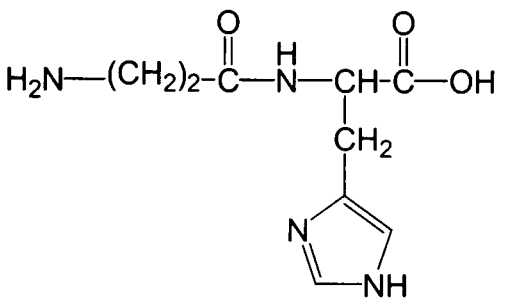
168



169

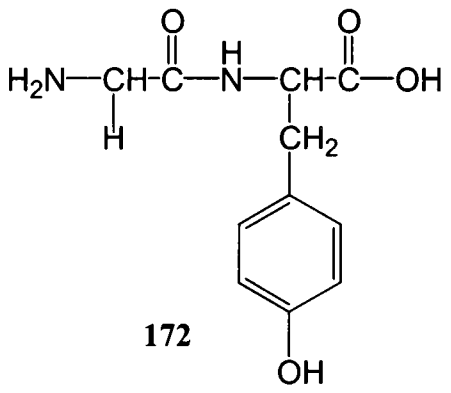


170

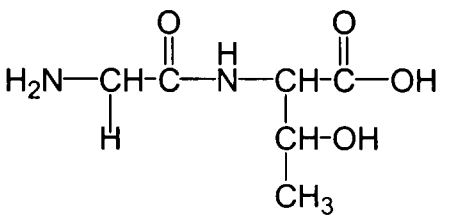


171

Dipeptides containing a hydroxyl residue

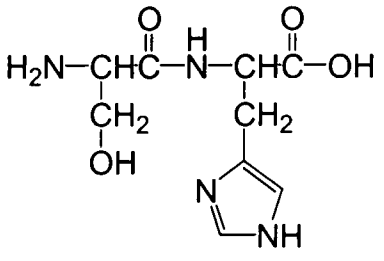


172



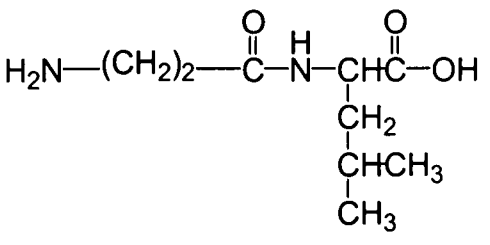
173

Dipeptides containing both basic and hydroxyl residues

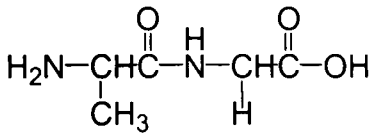


174

Dipeptides containing neither basic nor hydroxyl residues



176



175

Figure 24

4.1 Measurement of Translational Diffusion Coefficients Using PFG NMR Studies.

When two molecules form a complex together in solution, many NMR sensitive parameters change, such as chemical shift of the protons involved, T_1 and T_2 relaxation rates, and diffusion rates. Many examples have been reported in the literature, but most involve binding between a small ligand and a large receptor such as a protein[120]. Monitoring of small molecule-small molecule binding such as we were hoping to see is more complex since the changes observed are not as large.

Perhaps the most common method of determining binding has been to examine the change in chemical shift with varying concentrations of the two components. The signals most commonly used are amide NH peaks[121] (although other signals have been used[37]) since these peaks are downfield and usually distinct from the rest of the spectrum, and any change can be clearly observed. However, these changes are usually monitored in organic solvents such as chloroform. In our situation we wished to monitor binding in water, since the reaction to be catalysed is ester hydrolysis, and the binding between receptor and transition state has to occur in aqueous solution. This posed a problem for observing chemical shift change. Firstly, if deuterium oxide is used as the solvent, then the amide NH peak is no longer observed in the spectrum due to deuterium exchange. Water could be used as the solvent, but some of the amide resonances overlapped with the intense water peak at 4.6ppm, and if water suppression was applied then the overlapping peaks were also suppressed. In view of this, it was decided instead to observe changes in translational diffusion rates upon complexation. This is a technique which has largely been used to look at large-molecule, small-molecule binding[122,123], and we were therefore very interested to see if it could be successfully applied to our system in order to study binding between two small molecules.

The translational diffusion coefficient of a molecule in solution is molecular mass dependent: small molecules diffuse faster than large molecules. Therefore, when a ligand forms a complex with a host, the apparent molecular mass changes, and this can be observed as a slowing, or decrease in the translational diffusion coefficient. This change can be monitored using Pulsed Field Gradient (PFG) NMR techniques[124,125].

In NMR experiments, nuclear spins precess about the magnetic field at a frequency that is defined by their chemical identity and local electronic environment. Ignoring field inhomogeniety, all spins experience an identical magnetic field. The application of a field gradient has the effect of making the magnetic field strength linearly dependent on position.

When a gradient pulse is applied to the system the spins no longer have a coherent phase, and instead the phase of the individual spins becomes dependent on their transverse position. If no translational diffusion has occurred the spins can return to their coherent phase by application of a second gradient pulse of inverse polarity, and no loss of NMR signal will occur. However, if the spin has undergone translational diffusion, the second gradient pulse cannot realign the phases of the spins, and the resulting NMR signal will appear attenuated. The smaller the molecule, the further it will have diffused from its starting position, and the greater the attenuation of the signal. The intensity of the NMR signal can described by Equation 2[126]:

$$I = I_o \exp[-D(\Delta - \delta/3 - \tau/2)\gamma^2 g^2 \delta^2] \qquad \textbf{Equation 2}$$

Where I_o is the intensity of the resonance in the NMR spectrum in the absence of gradient pulses, γ is the gyromagnetic ratio, Δ is the diffusion delay time, g and δ are the amplitude and duration of the bipolar gradient pulse pair, D is the diffusion coefficient, and τ is the delay between the bipolar gradient pulse pair.

If the field gradient is varied, and all other NMR parameters kept constant, this can be simplified to Equation 3. A plot of LnI against the field gradient squared affords a straight line with the gradient of the slope proportional to the diffusion coefficient D.

$$\ln I = \ln I_o - CDg^2 \qquad \textbf{Equation 3}$$

Where $C = \left(\Delta - \delta/3 - \tau/2\right)\gamma^2 g^2 \delta^2$

The formation of the complex is an equilibrium process, and usually the exchange between free molecules and complex occurs faster than the NMR timescale. In this case,

the measured diffusion coefficient of, for instance the ligand, will be the mole fraction weighted average of the diffusion coefficient of bound and free molecules (Equation 4). In cases where the host is considerably larger than the ligand, it can be assumed that the diffusion coefficient of the complex (D_{HL}) is the same as that for the free host, and thus the percentage of bound ligand can be calculated from Equation 4. In our case however this assumption cannot be made, since the difference in molecular weights between the TSA and peptides was relatively small.

$$D_{obs} = X_L D_L + X_{HL} D_{HL} \qquad \textbf{Equation 4}$$

Where D_{obs} is the observed diffusion coefficient, D_L is the diffusion coefficient of free ligand, D_{HL} the diffusion coefficient of the complex, and X is the mole fraction.

Using this technique the translational diffusion coefficients of the peptides and TSA were measured both in free solution and in a mixture containing one peptide, and TSA in D_2O at pD7. Neutral pD was used, since the aim of the project was to synthesise a polymer that would catalyse ester hydrolysis under neutral conditions. Binding studies therefore had to be performed under similar conditions. The ratio of concentrations of TSA:peptide used was 30:3. The formation of the complex is an equilibrium process. Initially measurements used a 1:1 ratio, but very small changes in diffusion coefficient were observed. By Le Chatelier's Principle, use of an excess of the TSA would push the position of the equilibrium to the right, and encourage formation of any complex.

The diffusion coefficient of the TSA was measured to be 4.05×10^{-10} m^2/s in free solution. In the presence of any of the peptides, it was expected that this value would not significantly change, since even if a very stable complex formed only 10% of the TSA molecules could be involved in binding at any one time. We were therefore most interested in observing the change in diffusion coefficients of the peptides. Peptides showing the strongest binding should show the greatest decrease in diffusion coefficient. The results are summarised in Table 1. The first two columns show the measured diffusion coefficients. The third column is the change in diffusion coefficient as percentage value. The values are accurate to $\pm 0.1 \times 10^{-10} m^2/s$.

Peptide	D_{pep} (x $10^{-10}m^2/s$)	D_{obs} (x $10^{-10}m^2/s$)	$\Delta D/D_{pep}$ x 100 (%)
H-Arg-Arg-OH	5.01	4.06	18
H-Ala-Arg-OH	5.62	5.01	11
H-βAla-Lys-OH	5.82	5.30	9
L-Carnosine	6.20	5.69	8
H-Ala-Gly-OH	7.22	6.79	6
H-Gly-Tyr-OH	5.90	5.56	5
H-β-Ala-Leu-OH	6.07	5.81	5
H-Ser-His-OH	5.91	5.70	3.5
H-Gly-Thr-OH	6.67	6.51	3

Table 1

All of the dipeptides are unprotected, and therefore contain at least one free amino group at their N-terminus. Therefore all should be capable of forming a complex, however weak, between the amine and the carboxylic acid group on the TSA. It was expected in all cases to observe a slight change in the diffusion coefficient upon mixing with the TSA. As can be seen from Table 1, this was indeed the case.

Using the information obtained from the Cambridge Structural Database, it was expected that the dipeptides containing basic, or hydroxyl residues should form the strongest complexes with the TSA, since the terminal amino group could interact with the carboxylic acid group, and the hydroxyl or basic residue interact with the phosphonic acid. As shown in Table 1, the four peptides containing one or more basic residues show the largest change in diffusion coefficient upon mixing. Surprisingly the three peptides containing hydroxyl residues did not appear to show good affinity for our

TSA. This may have been due to conformational restrictions in the case of the tyrosine residue, or steric hindrance caused by branching in the dipeptides containing serine and threonine.

Of the four peptides containing basic residues, L-Carnosine **171** showed the weakest complexation to the TSA. Two possible explanations can be put forward. Firstly, at pD7, we would expect the imidazole ring to be neutral, since the pKa of the imidazole ring in histidine is 6.0. Secondly, the rigidity of the 5-membered ring may prevent the molecule achieving an energetically favourable conformation for binding. Of the other three basic dipeptides, the two that showed the greatest change were Ala-Arg **169**, and Arg-Arg **168**. This comes as no real surprise, since the side chain of arginine contains a guanidine group, which is known to form stable complexes with carboxylate groups, as shown in Figure 25[127]. It is highly possible that a phosphonic acid might be able to interact in similar manner.

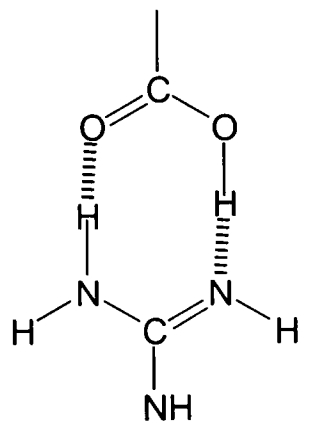


Figure 25

Of the two dipeptides that contained neither a basic nor a hydroxyl residue, one, β-Ala-Leu **176** showed only a very small change in diffusion coefficient as expected. However, the other peptide, Ala-Gly **175**, although not showing a change as large as that observed with the basic dipeptides, still showed a significant decrease in the value of D_{pep}. Possibly due to the simple structure of the molecule, it can find a favourable conformation in which the terminal carboxylate is able to interact with the phosphonic acid on the TSA.

It is worth noting at this point that the changes observed in the diffusion coefficient of these dipeptides are only an indication of binding. As mentioned above, due to the small

size of our two molecules, it is not possible to determine a binding constant from the data collected. However, the results are enough to clearly indicate which peptides show complexation to the TSA. As the results seemed to show that Arg-Arg **168** was the best binder, and showed the most potential to be incorporated into our polymer catalyst, we decided to measure the binding constant K_a for the complex between **168** and the TSA **146**.

4.2 Measurement of the Binding Constant for the Complex between Arg-Arg and TSA.

The binding constant for a bimolecular host-ligand complex (Equation 5) can be given by Equation 6[128].

$$TSA + Pep \underset{}{\overset{K_a}{\rightleftharpoons}} Complex \qquad \textbf{Equation 5}$$

$$K_a = \frac{[Complex]}{[TSA][Pep]} \qquad \textbf{Equation 6}$$

Equation 6 can be used to relate the binding constant to the measured diffusion coefficient. In a situation where the equilibrium process is occurring faster than the NMR timescale, the observed diffusion coefficient is the mole fraction weighted average of the diffusion coefficients observed in the free and complexed molecule (Equation 4).

$$D_{obs} = X_L D_L + X_{HL} D_{HL} \qquad \textbf{Equation 4}$$

In systems where the host is considerably larger than the ligand, the diffusion coefficient of the host should not be affected upon complexation, and hence D_{HL} is assumed to be equal to the diffusion coefficient of the free host. The mole fractions can therefore be found from Equation 4, and the binding constant K_a calculated. In cases such as our system, where the two molecules are similar in size, no such assumption can be made, and there are therefore two unknown parameters: K_a and D_{HL}. These two parameters can be calculated by measuring the diffusion coefficients with a series of

differing concentrations of ligand and host, and using a non-linear curve fitting procedure[128].

The principle of curve fitting methods is that if the binding stoichiometry is known, a binding isotherm may be calculated and compared to the experimental data. D_{HL} and K_a are separate variables and the correct values of these two variables are those that produce the best fit of calculated to observed data.

The concentration of Arg-Arg **168** was kept constant (1mM), and the concentration of the TSA **146** was varied from 0-52mM. The diffusion coefficient was measured at all concentrations using the method described above. The results are shown in Table 2.

Ratio Peptide : TSA	**D_{pep}**
1 : 0	**5.17**
1 : 1	**4.82**
1 : 1.8	**4.74**
1 : 3.9	**4.57**
1 : 6.8	**4.46**
1 : 8.7	**4.40**
1 : 15.8	**4.35**
1 : 18.6	**4.25**
1 : 27.4	**4.20**
1 : 37	**4.15**
1 : 52	**4.11**

Table 2

The diffusion coefficient of the peptide in the complex was plotted against the concentration of the TSA, and the binding constant K_a was calculated using the non-

linear least squares curve-fitting programme Associate V15[#129,130]. The calculated value was found to be $393LM^{-1}\pm10\%$.

Weak binding generally has a value of K_a less than$10LM^{-1}$, and strong binding a value of K_a greater than 10^5LM^{-1} [128]. Our complex has a binding constant in the middle of these two values, and therefore it can be deduced that the peptide has a reasonably strong affinity for the TSA. It would not have been expected to see strong binding in a system such as this, since the use of water as a solvent interferes with much of the hydrogen bonding between, for instance, the guanidine group and the carboxylate. Should the studies have been performed in a non-polar aprotic solvent such as chloroform, we would have expected to observe a much higher value for K_a.

Although these data told us that the Arg-Arg **168** was binding to the TSA, it was only possible to estimate how, and whether it was acting as a bidentate ligand as envisaged (Figure 23). Molecular modelling studies were therefore performed to learn more about the possible structure and conformation of the actual complex.

4.3 Molecular Modelling Studies.

The aim of the molecular modelling studies was to look at possible interactions between the TSA **146** and dipeptide Arg-Arg **168** in aqueous solution. The intended method was to do this by performing conformational analysis on the TSA and dipeptide to find low energy conformers that matched known geometrical constraints. These conformers could then be docked together and the energy minimised to find low energy complexes.

Both the TSA **146** and dipeptide **168** were built in Sybyl[131]. The calculations were performed as if at pH7, in which case both guanidine groups, and the terminal amino group of the dipeptide would be protonated, and the carboxylate and phosphonic acid groups of TSA **146** deprotonated. All the energy calculations were performed using the Tripos force field. A distance dependant dielectric constant of 80 was used in order to take account of any solvent interactions[132]. Water plays an important role in determining conformational preferences of polar molecules, and in modulating the forces between

[#] Programme generously donated by Dr. B. R. Peterson (See Ref 130.)

ligand and host. Accounting for solvation effects is essential for a qualitative understanding of the interactions between molecules. Four main methods have been developed to model the effects of solvents[133]. Due to time limitations, and the software available to us, we decided to use the distance dependent dielectric approach.

The force felt between two charges is generally accounted for by force fields using Coulomb's Law (Equation 7).

$$Vij = \frac{qiqj}{4\Pi\varepsilon rij} \qquad \textbf{Equation 7}$$

Where V_{ij} is the force felt between two charges, q_i and q_j, r_{ij} is the distance separating the two charges and ε is the dielectric constant of the medium. If the dielectric constant in the molecular mechanics simulation is increased from the value for vacuum (ε=1) to that of water (ε=80) then the electrostatic screening effect of the solvent may be simulated. This treatment, as used by some force fields can however overestimate the forces felt between molecules at long distance, as the charges are interacting through bulk solvent and are more heavily screened. The Tripos force field employs a distance dependent dielectric, where the dielectric constant in Coulomb's Law is replaced by d_{rij} (d= a proportionality constant) and the force now decays as $1/r_{ij}^2$. This method reduces the large forces experienced at long distances, but does not completely eliminate them.

Both molecules were energy minimised using the Tripos force field and a distance dependant dielectric function of 80. The TSA was then subjected to a conformational analysis. Conformational analysis can be performed in a number of ways[134]. The simplest method is energy minimisation, in which the atoms of a molecule are moved in such a way is to reduce the total energy of the system based on an empirical representation for the interaction energy of the atoms of a molecule. However, when this procedure is carried out from the same starting point it always produces one single conformation, and therefore cannot guarantee to find the conformation with the global minimum energy. At the other extreme is systematic conformational analysis, in which a set of rotatable bonds is identified. All possible rotamers of each of these bonds are examined as a function of each other. This method identifies a full set of conformations, but if the molecule contains a very large number of rotatable bonds, is a very time

consuming method producing vast numbers of conformations. We therefore elected to use the Random Search method[135].

Random Search is a conformational analysis method in the Sybyl package that combines random torsional perturbations with energy calculations, minimisation and conformer comparison in an iterative process in order to output sets of unique conformations. The measure of completeness is based on recording how often each conformer is found, and the random search will find the same conformations many times. If every conformation has been found at least n times there is a (1-(1/2)n probability that all possible conformations have been found. Thus if each conformation has been found five times there is a 96.9% chance that all possible conformations have been found.

When a Random Search was performed on the TSA using the Tripos force field, a database of over 1500 possible conformations was generated. The database was sorted in order of increasing energy, and the lowest energy conformer subjected to minimisation. This gave a global minimum energy conformation of the TSA of 34.373 kcal/mol. This was a huge number of possible conformations. The possible conformations were reduced in number by measuring some geometrical constraints, and then applying them to the library of possible conformers.

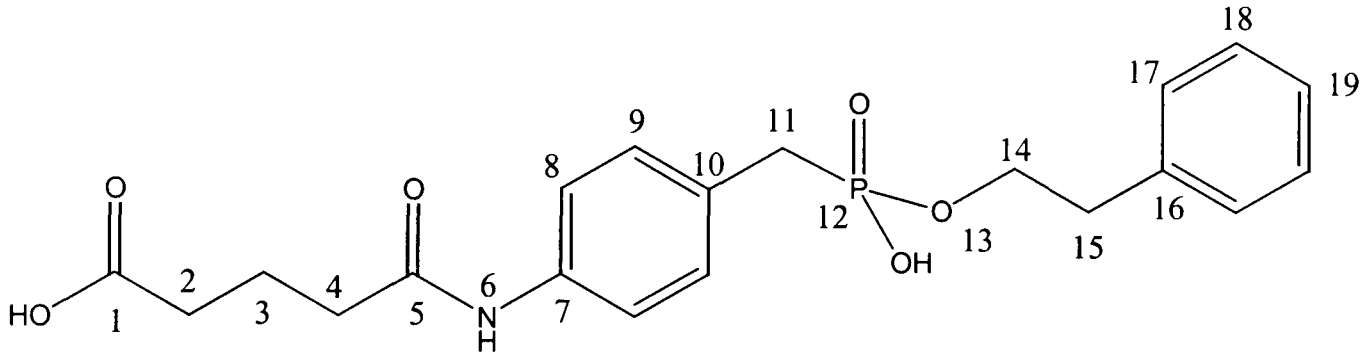
Internuclear distances of the TSA molecule were measured using 2D-NOESY NMR spectroscopy. The internuclear distance between two spins (r_{ab}) can be calculated by comparing the volume integral intensity of the cross peak ($intensity_{ab}$) in the NOESY experiment with that ($intensity_{ref}$) for a pair of protons of known internuclear separation r_{ref} (Equation 8)[136].

$$r_{ab} = r_{ref}\left[\frac{\text{intensity}_{ab}}{\text{intensity}_{ref}}\right]^{-\frac{1}{6}} \qquad \textbf{Equation 8}$$

In our system the TSA contains a *para*-substituted phenyl ring. The distance between *ortho* protons in a system such as this is known to be 2.8Å[136]. Therefore, if the intensity

of all the cross peaks were measured and Equation 8 applied, we should have some idea of internuclear distances in the TSA.

2D-NOESY NMR spectra of TSA **146** were obtained in D_2O at pD7 at a variety of mixing times ranging from 100-600ms. The maximum cross-peak intensities were found at 300ms. Using the knowledge that the distance between two *ortho* protons in a *para*-substituted phenyl ring is approximately 2.8Å, the volume integral intensities of all the cross peaks in the spectra were measured, and inserted into Equation 8 to calculate internuclear distances. The results are shown in Table 3.

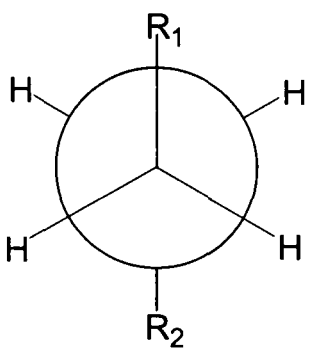


Geometrical Feature	**Constraint**
Distance 8H-9H	2.8Å
Distance 9H-14H	3.62-3.75 Å
Distance 14H-17H	3.26-3.62 Å
Torsion Angle 1, 2, 3, 4	180°
Torsion Angle 2, 3, 4, 5	180°
Torsion Angle 13, 14, 15, 16	180°

Table 3

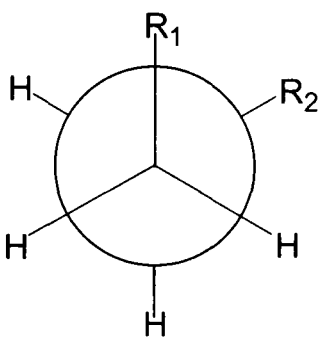
As can be seen from Table 3 some torsion angles in the TSA were also measured. This was done simply using the ^1H NMR spectra of **146**. The CH_2 at positions 2, 4, 14 and 15 all appear as triplets. Interpretation of this data means that the two protons of the CH_2

are chemically equivalent, but magnetically inequivalent. The only conformation that can fit this scenario is when the torsion angle is 180° as shown by **177** in Figure 25. If the torsion angle is as that shown in **178**, then the two protons are now in different chemical environments, and a multiplet would be observed. Although there is free rotation around all the bonds, there is almost 100% population of this "trans" conformation.



177

Torsion angle 180°



178

Protons no longer chemically equivalent

Figure 26

The constraints in Table 3 were then applied to the library of conformations for **146** calculated during the Random Search. Four conformations were found to match the geometric constraints. Two of these conformations however showed distinct distortion of one of the phenyl rings, and were therefore discarded from use in further studies. This left two conformers which we could use to dock with Arg-Arg.

A similar Random Search was performed on the dipeptide, Arg-Arg **168**. Again the conformers were sorted by energy in order to find the global minimum of 13.253Kcal/mol. In order to find conformers suitable for docking with the TSA **146**, the distance between the carboxylate and phosphonic acid groups of the two TSA conformers was measured and found to be 12-14 Å. Only conformers of **168** in which the distance between either the two guanidine groups, or the terminal amino group and one of the guanidine groups, was similar to this value were considered for docking against the TSA. This left a set of 10 possible conformers.

Different conformations of the dipeptide **168** and TSA **146** were docked together by eye. Out of the all the possible combinations, four complexes, Temp137Arg618, Temp137Arg193, Temp137Arg894, Temp137Arg1760, were studied further, as they showed the potential for hydrogen bonding between both a guanidine group and the phosphonic acid, and a guanidine group and the carboxylate. All four of these involved the same TSA conformation. Using the DOCK command, the dipeptide was fitted interactively to the TSA in such a way as to maximise the electrostatic interactions, and minimise the steric energy.

The first attempt to minimise the energy of the complex used the MINIMISE_DOCK command. However, this tool concentrates only on minimising the electrostatic energy, and consequently significantly altered the conformations of both TSA and peptide such that the energies of the individual components were now very unfavourable. Instead the entire complex was minimised under three different sets of conditions: 1) Tripos forcefield, distance dependent dielectric of 80; 2) Tripos force field, distance dependent dielectric of 1; MMFF94s force field, dielectric constant of 1. The minimised energies of the complex, and of the two components were measured, and used to calculate the interaction energy of the complex using Equation 9:

$$E_i = E_{total} - (E_{pep} + E_{tsa})$$ **Equation 9**

Where E_i is the interaction energy, E_{total} the energy of the complex, and E_{pep} and E_{tsa} the energies of the individual components. The results are summarised in Tables 4-6.

1) Tripos force field, distance dependent dielectric 80

Complex	Energy of Complex	Energy of TSA	Energy of Peptide	Interaction Energy
Temp137Arg618	8.231	3.137	16.098	-11.004
Temp137Arg193	8.000	3.324	17.554	-12.878
Temp137Arg894	11.053	2.992	18.253	-10.193
Temp137Arg1760	9.913	3.593	17.610	-11.29

Table 4

2) Tripos force field, distance dependent dielectric 1

Complex	Energy of Complex	Energy of TSA	Energy of Peptide	Interaction Energy
Temp137Arg618	-162.398	-8.838	5.928	-160.108
Temp137Arg193	-132.286	-13.741	-17.623	-100.922
Temp137Arg894	-128.668	-14.610	-31.405	-82.653
Temp137Arg1760	-139.555	-10.720	-28.667	-100.168

Table 5

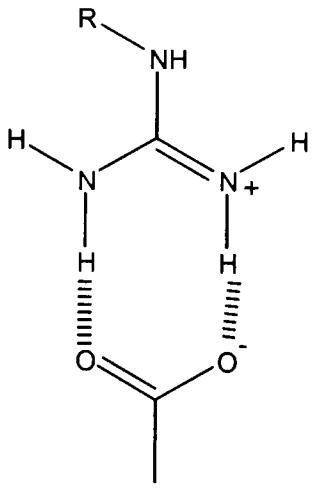
3) MMFF94s force field, dielectric constant 1

Complex	Energy of Complex	Energy of TSA	Energy of Peptide	Interaction Energy
Temp137Arg618	-641.665	84.525	-325.537	-400.653
Temp137Arg193	-610.686	73.343	-356.559	-327.47
Temp137Arg894	-528.367	65.130	-332.247	-261.25
Temp137Arg1760	-634.320	74.188	-342.666	-365.842

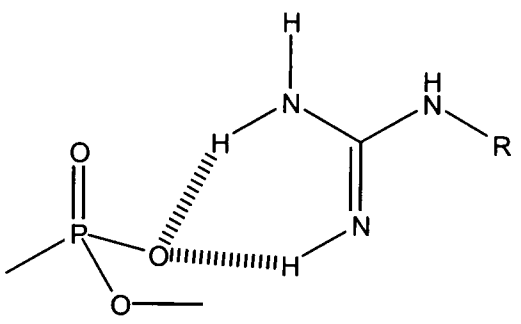
Table 6

As can be seen from Tables 4-6 all of the complexes formed had favourable interaction energies. Although the minimisations with Tripos (distance dependent dielectric 1) and the MMFF94s force fields gave complexes with favourable interaction energies, when the complexes were studied in further detail, it was clear that the conformations of **146** and **168** had been distorted in order to maximise the electrostatic interactions. Therefore the complexes obtained using the Tripos field, distance dependent dielectric 80 were used for further study. The complexes were then compared to known small molecule close contacts.

Isostar is a computerised library containing information about the non-bonded contacts formed by some 250 chemical groupings[137,138,139]. This database was used to study the geometrical preferences of interactions between guanidine groups and both carboxylate and phosphonate groups. The results are shown in Figure 27.



Guanidine-carboxylate interactions



Guanidine-phosphonate interactions

Figure 27

This information was compared with the geometries observed in our complexes. We were pleased to note that in two of the complexes, Temp137Arg618, and Temp137Arg193, the geometries observed closely matched those in the database, although the guanidine-carboxylate interaction is usually more planar than that observed in our complex. Three different diagrammatic representations of the complex Temp137Arg618 are shown in Figures 28-30.

The combination of NOE NMR spectroscopy and molecular modelling has given us a clear diagrammatic representation of how the TSA molecule **146** and dipeptide Arg-Arg **168** could complex together. As envisaged, it does indeed seem possible for the two guanidine moieties to interact with the carboxylate and phosphonate on **146** as a bidentate ligand. In the modelled complex, it is also pleasing to note that the carboxylic acid group of the dipeptide is not involved in any binding, and can therefore be utilised to attach the group to a polymer backbone. The only slight disappointment, as can be

seen in Figure 28, is that the hydrogen bond which we had envisaged between the two amide groups is not observed. In addition, when the complex is compared with known small molecule complexes, it can be seen the geometries of the interactions in our complex are very similar to that observed in actual crystallographic data. The information obtained from the modelling studies, and the diffusion rate measurements indicated that Arg-Arg should be able to act as a receptor site capable of binding the transition state of our reaction. It was therefore decided to incorporate this moiety into our potential catalytic polymer.

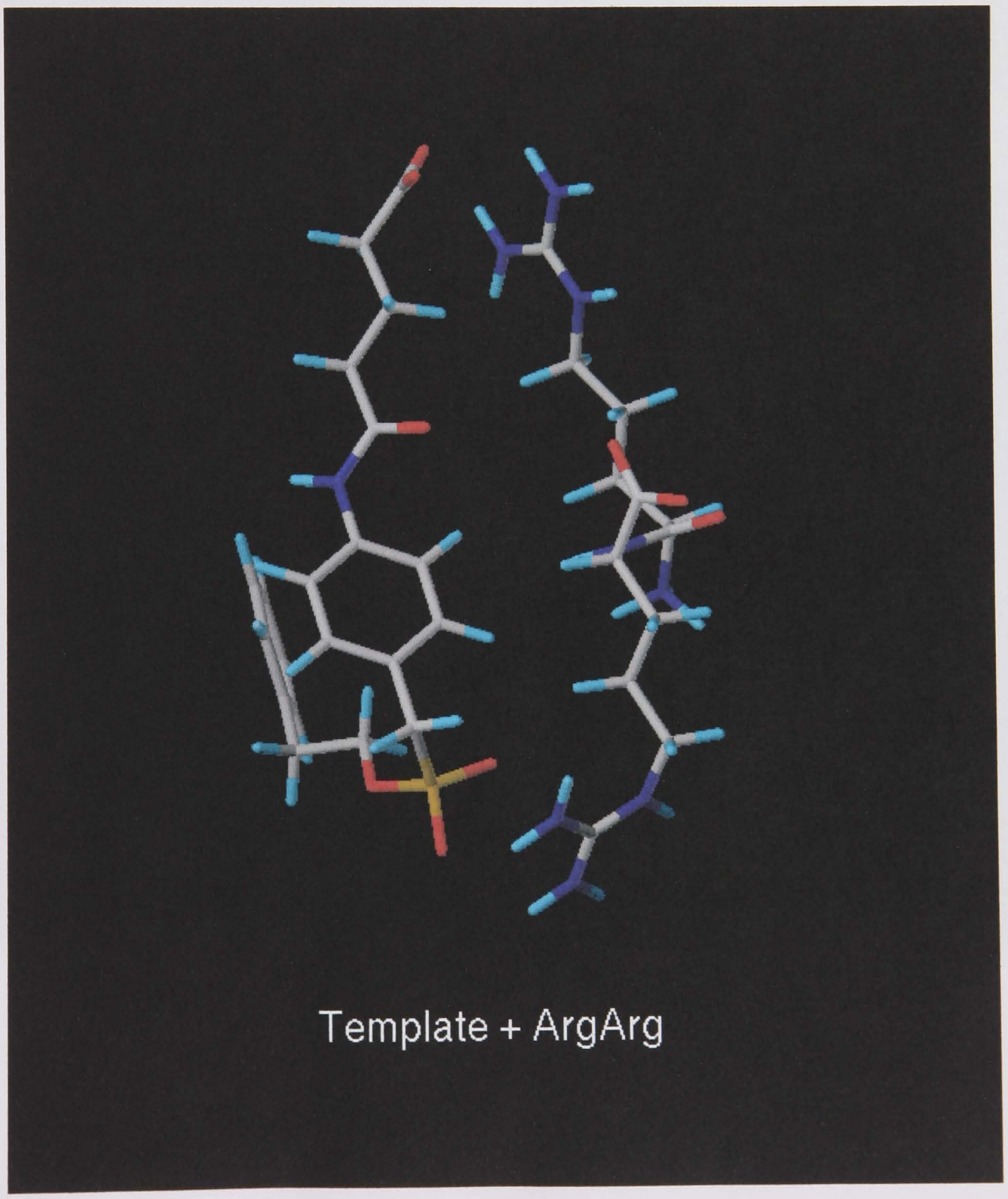


Figure 28

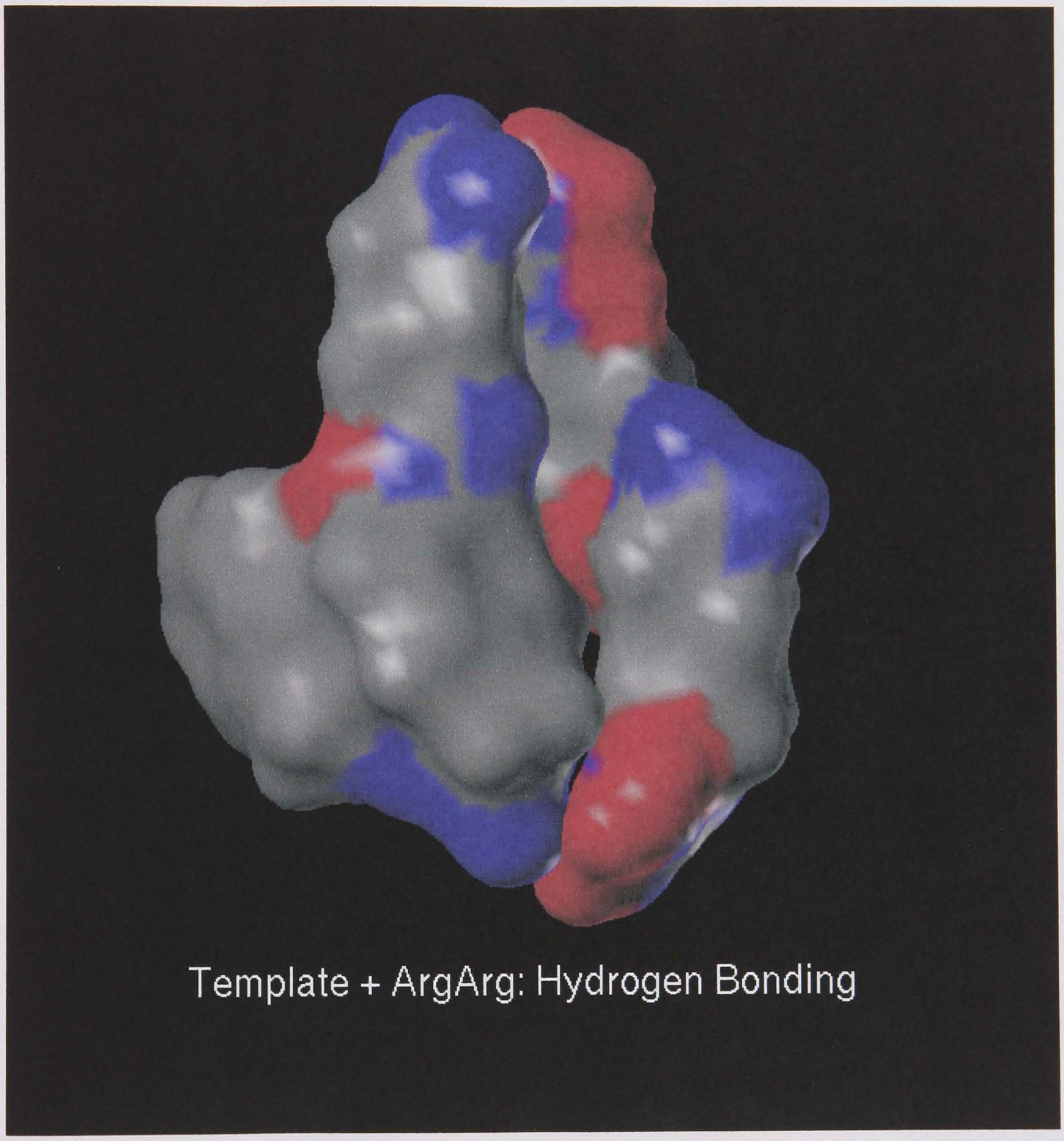


Figure 29

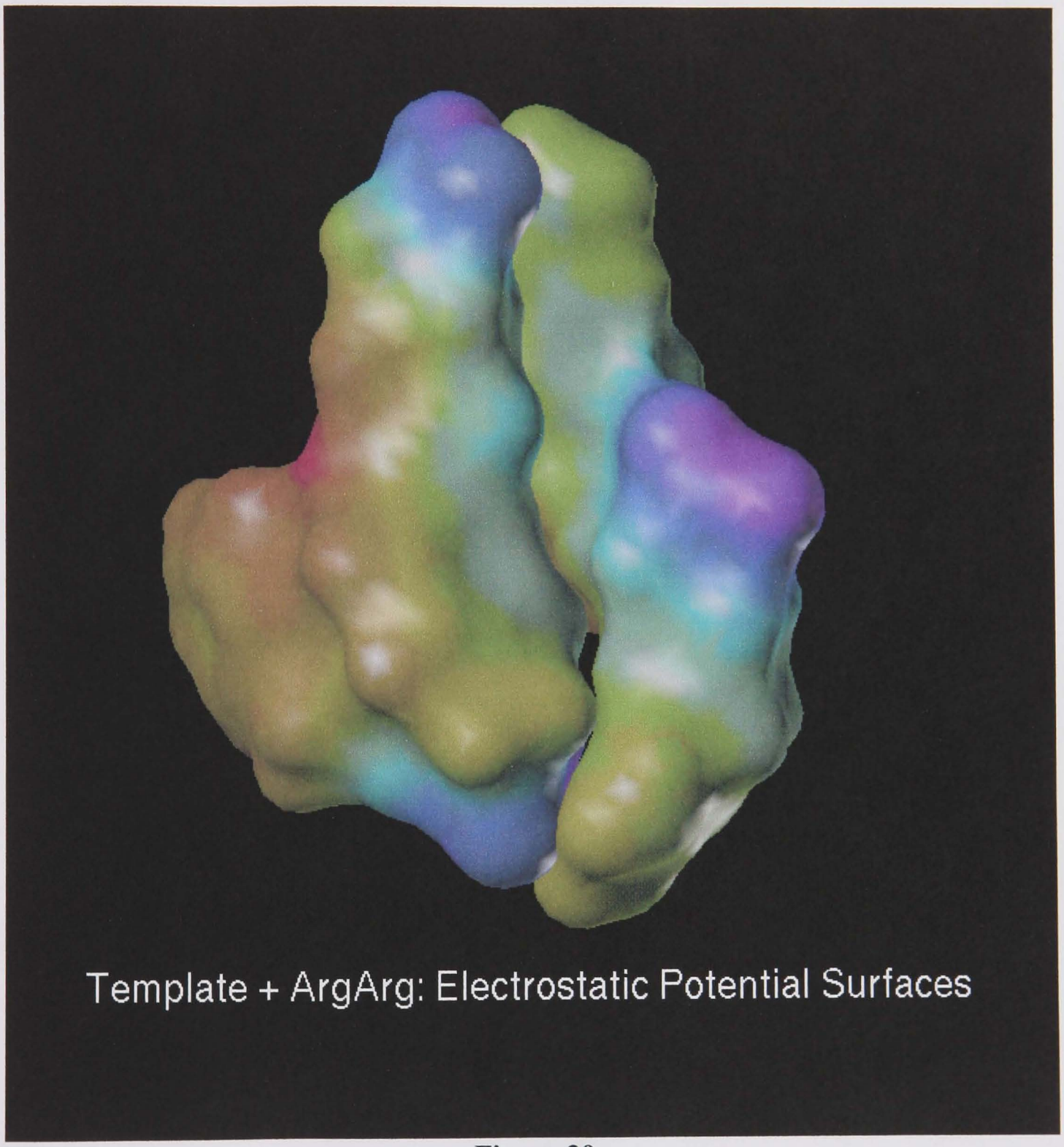


Figure 30

5. Development of the Polymer Catalyst.

With the dipeptide Arg-Arg **168** chosen as the "receptor site" the next step was to decide on which functional groups to incorporate into the polymer to act as "catalytic groups". As the receptor site was comprised of a dipeptide, it seemed both logical and "natural" to use amino acid residues as our active functional groups as well. The carboxylic acid moiety of the amino acid could of course be utilised to attach the group onto the polymer backbone.

With this in mind, data on the active site of the proteases, was examined. Proteases can be split into four classes, according to their activities and functional groups[140]. The serine proteases have a reactive serine residue and operate at neutral pH. The thiol proteases are similar to serine proteases, but with an active cysteine residue replacing the serine residue. Their optimum pH is also around 7. The carboxyl proteases contain catalytically important carboxyl (e.g. aspartic acid) residues, and tend to operate at low pH. Finally, the zinc proteases are metalloenzymes which also operate at neutral pH.

Since the objective of the project was to create an artificial enzyme which could catalyse ester hydrolysis at neutral pH, the carboxyl proteases were therefore disregarded, and attention focused on the other three classes. The mechanism of the serine proteases has been discussed in detail in Part 1, with the conclusion that catalysis is a result of a triad of residues, His-57, Ser-195 and Asp-102, as shown in Figure 31[5,6,7]. The key step of the reaction relies on hydrogen transfer from the hydroxyl group of the serine residue to the imidazole ring of the histidine. Thus the hydroxyl can then act as a nucleophile and attack the carbonyl group of the substrate to form a tetrahedral intermediate. It was therefore logical to investigate the use of both serine, and histidine as catalytic functional groups in an artificial enzyme.

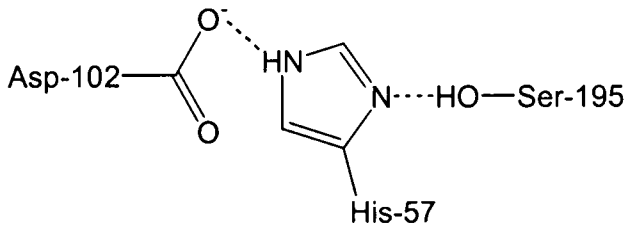
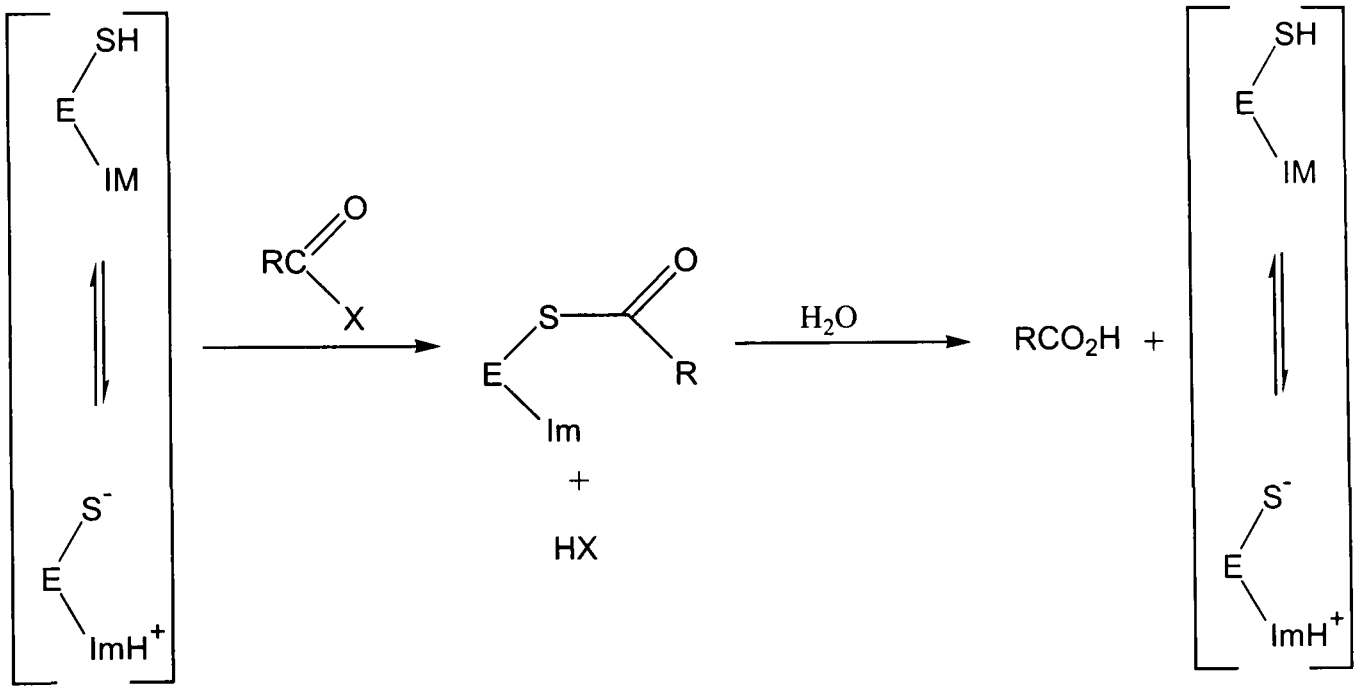


Figure 31

Thiol proteases such as papain, ficin and bromelain are widely distributed in nature[140]. In a similar manner to serine proteases, a cysteine and histidine residue act in concert. Initially it was believed that the mechanism was virtually identical to that of the serine proteases, i.e. histidine general base-catalysed attack of the cysteine on the substrate to form an acyl enzyme followed by deacylation. However, there was no evidence from model systems that thiol nucleophiles can be activated by general base catalysis. Instead it is believed that the active form is [$RS^-.ImH^+$], where Im represents the histidine. The currently accepted mechanism is shown in Scheme 49. Using this information, cysteine was also chosen for incorporation into the polymer.



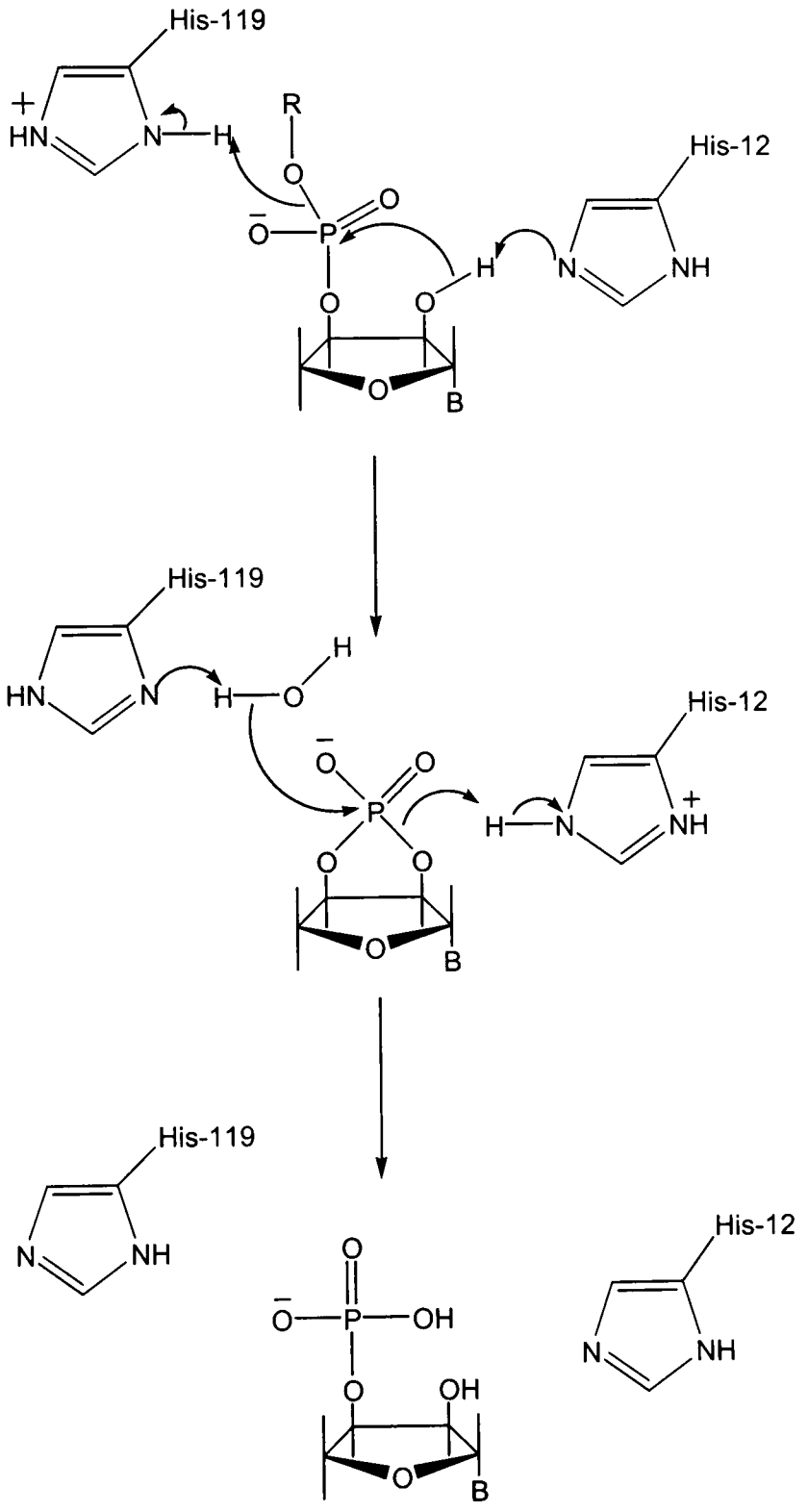
E = Enzyme, S = Sulfur of cysteine, Im = Imidazole of histidine

Scheme 49

The third class of protease, zinc proteases, were also disregarded at this stage since none of the modelling or binding studies had been performed in the presence of a metal ion, and it was uncertain how it would affect the mode of binding of the substrate to the "receptor unit". If required, the mode of action could be re-examined at a later date.

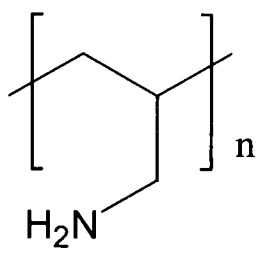
With histidine chosen as one of the functional groups, it is also worth discussing the action of Ribonucleases, since they demonstrate how two histidine residues can function together as a general acid/base pair[140]. Bovine pancreatic ribonuclease catalyses phosphate hydrolysis by a two-step process in which a cyclic phosphate intermediate is

formed (Scheme 50). The reaction is catalysed by concerted general-acid-general base catalysis by the His-12 and His-119 residues. In the initial cyclisation step, His-119 acts as a general acid to protonate the leaving group. In the second hydrolysis step, His-119 activates the attack of water by general-base catalysis, and His-12 protonates the leaving group. It was envisaged that in the microenvironment created by a polymer, it may be possible for two histidine residues to act in a similar manner to hydrolyse an ester substrate.



Scheme 50

We now had three potential catalytic residues for incorporation into our polymer. The next step was to choose the polymer backbone. The requirements were that it had to be commercially available, have functionality available to attach the peptide and amino acid residues onto the backbone, and preferably be soluble in a variety of solvents. With this in mind, polyallylamine **134** (Figure 32) was chosen. The primary amino groups are suitable for attachment of the amino acid residues by standard peptide coupling techniques. Although the commercially available form of polyallylamine is as its hydrochloride salt, it can be readily converted to the free base by stirring with ion-exchange resin, and the free base is soluble in a range of solvents, including water, chloroform, methanol and ethanol.



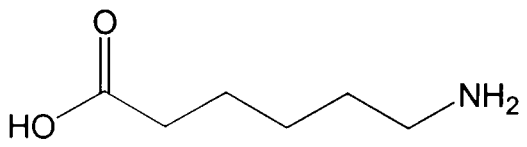
134

Figure 32

If standard peptide coupling techniques were to be used to attach the amino acids and peptide to the polymer backbone, then the side chains and α-amino groups had to be protected. It was decided to protect the α-amino groups of all the amino acids as their base labile Fmoc derivatives. The imidazole ring of histidine, and thiol group of cysteine were to be protected as trityl derivatives, the guanidine group of arginine as the Pbf derivative, and the hydroxyl group of serine as the tertiary butyl derivative. All of the side chain protecting groups chosen were acid labile.

It was important that the Arg-Arg dipeptide should be conformationally mobile, such that even when attached to the polymer backbone it could achieve the energetically favourable conformation required to bind to the transition state of the reaction. It therefore seemed logical to place a spacer between the polymer backbone and the peptide. Lysine initially seemed a suitable choice for the spacer, since it could be

coupled onto the polymer *via* the α-carboxyl group. The side chain amino group, which would be five carbons away from the polymer, could be coupled to the dipeptide. However, not wishing there to be complications arising from the peptide accidentally being coupled to the α-amino group, an analogue of lysine, 6-aminocaproic acid **179** (Figure 33) was used as the spacer molecule instead. To ensure that the arginine dipeptide was coupled to this linker, the pseudo "tripeptide" would be synthesised and attached to the polymer as a single unit.



179

Figure 33

At the same time it was decided to functionalise any unreacted amino groups from the polymer backbone with a lysine residue. The thought was that the amino groups of the lysine may be suitably positioned to aid with the catalysis, either as functional groups, or they may aid in the binding of the substrate. Although not often present in the active site of proteases, lysine residues do play a crucial role in the catalytic action of some enzymes (e.g. aldolases, see page 43). In the neutral environment of the polymer, the lysine residues should be protonated, and therefore could behave as a general acid catalyst.

The final part of the strategy was to decide upon the ratio of receptor:functional group incorporated into the polymer. Initial studies would involve the use of only one of the potential catalytic groups in each polymer. After the polymer had been tested for its catalytic capabilities, data could then be compared and combined such that the most effective groups could be incorporated together to form a "second generation" more efficient catalyst. A 1:1 ratio of arginine dipeptide: functional group was initially used, each occupying a third of the amino groups on the polymer backbone. The final third would be functionalised with lysine, to form a polymer such as that shown in Figure 34.

The catalytic residue and receptor dipeptide would be incorporated into the polymer at the same time such that they could influence the positioning of the other on the polymer backbone. As explained in section 1, each binding site will be surrounded by a number of catalytic groups, and only one needs to be in the right position and orientation relative to the substrate for the hydrolysis reaction to occur.

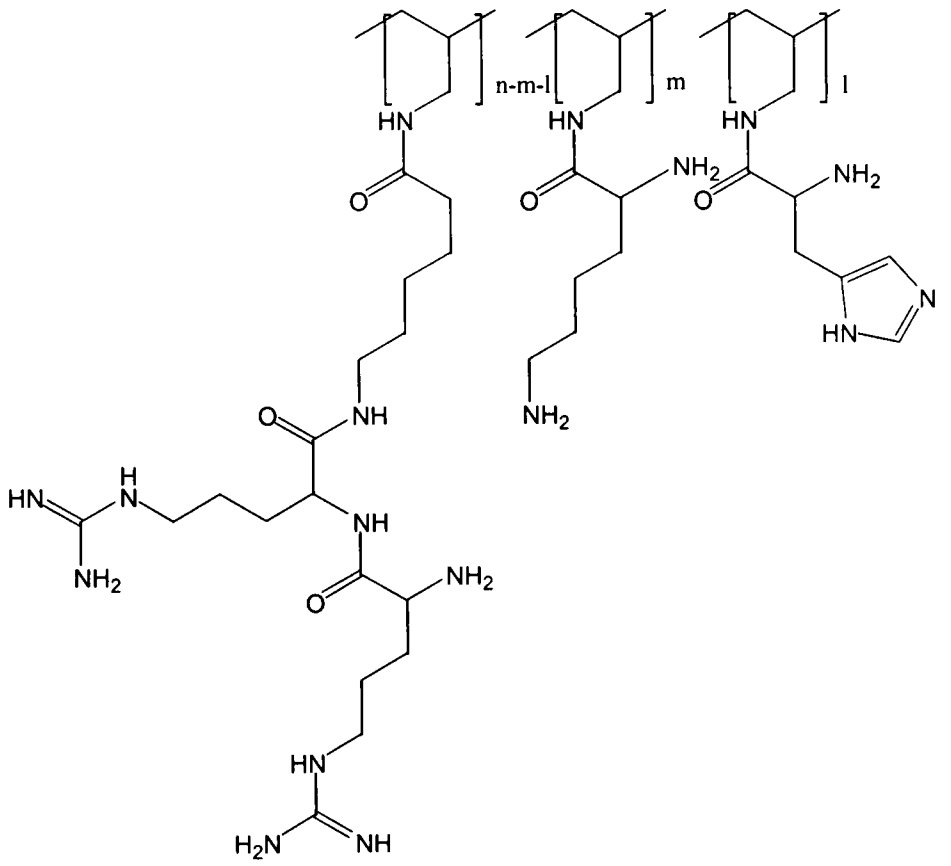
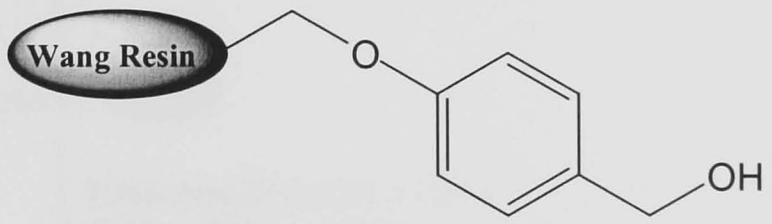


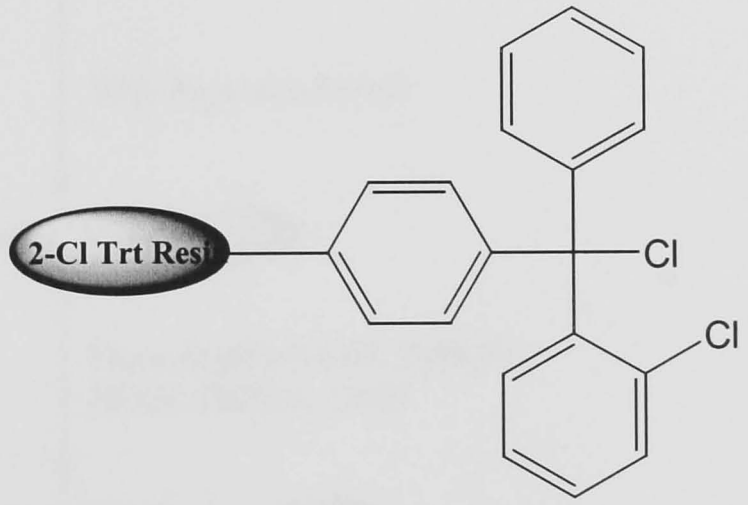
Figure 34

6. Synthesis of the Receptor Unit.

The receptor unit, consisting of the protected dipeptide attached to a 6-aminocaproic acid linker, **179**, was synthesised manually on the solid phase using standard Fmoc methodology. The acid labile 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) group was used to protect the guanidine side chain of the arginine residues[141]. Although at the end of the synthesis we required all protecting groups still to be in place, some level of orthogonality was preferred, so that when the peptide was attached to the polymer backbone side chains could be selectively deprotected if required.



180

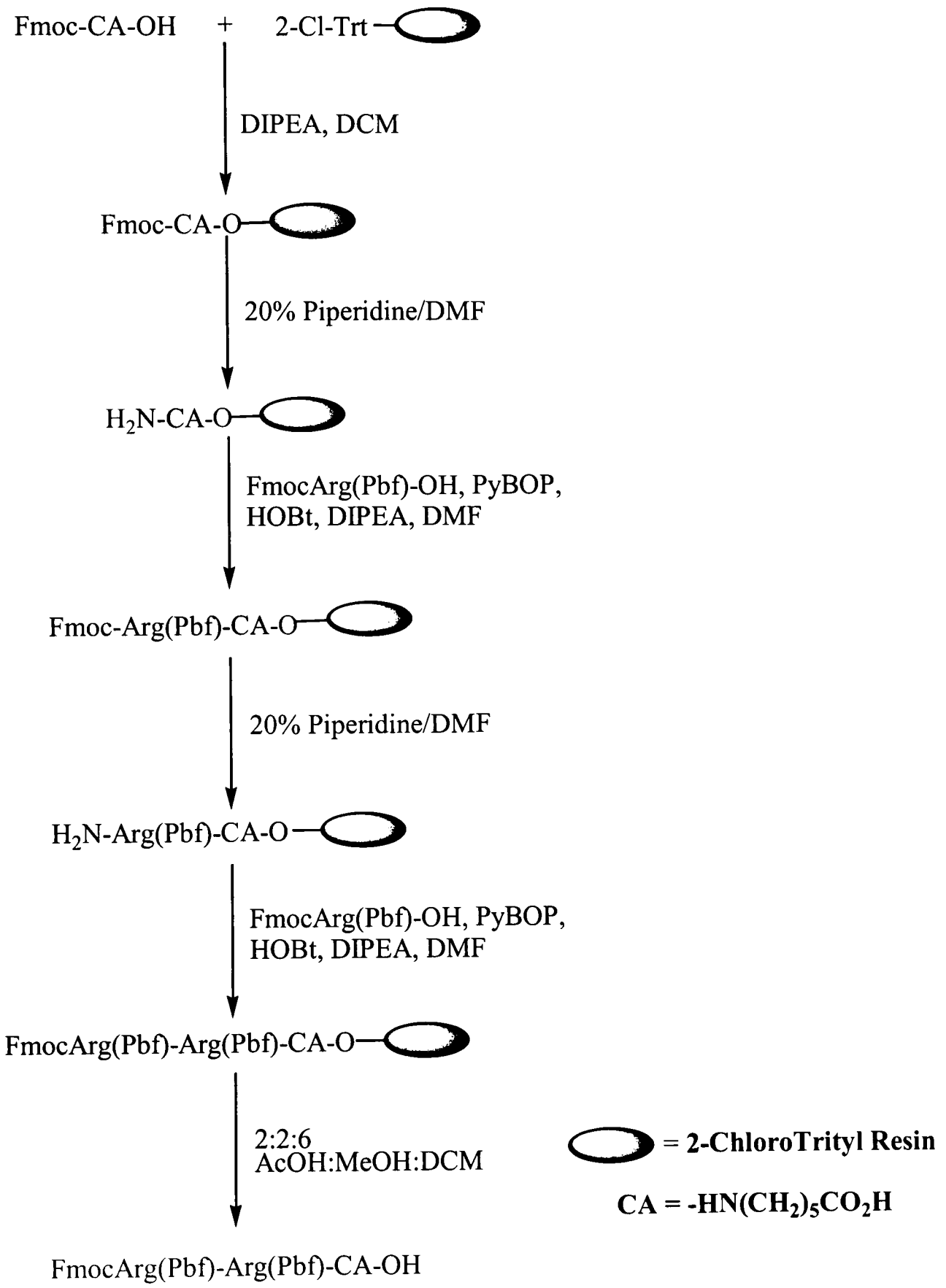


181

Figure 35

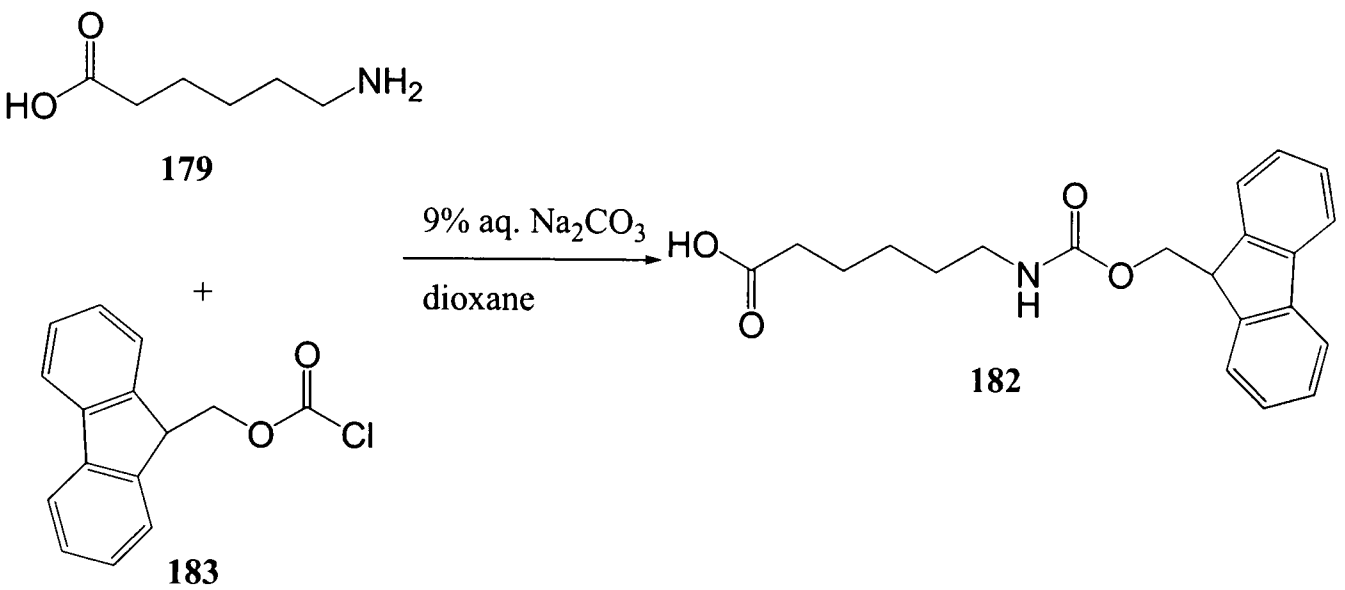
As it was a requirement of the synthesis that the peptide had to be fully protected after cleavage from the solid support, the choice of resin was crucial. Cleavage of the peptide from the chosen resin had to take place under mild conditions. The most commonly used resin is the Wang resin **180**[142] (Figure 35), but cleavage from this resin requires the use of TFA which would also cleave the Pbf protecting groups. Instead it was decided to use the 2-chlorotrityl resin **181**[143], since cleavage requires use of dilute acetic acid,

which should leave all the protecting groups in place. The procedure for the synthesis is outlined in Scheme 51.



Scheme 51

Fmoc protected 6-aminocaproic acid (CA) **182** was prepared according to literature procedure by adding 9-fluorenylmethyl chloroformate (Fmoc-Cl) **183** in dioxane to a solution of **179** in aqueous sodium carbonate (Scheme 52)[144]. The product **182** (Fmoc-CA) was obtained in an 86% yield following chromatography.



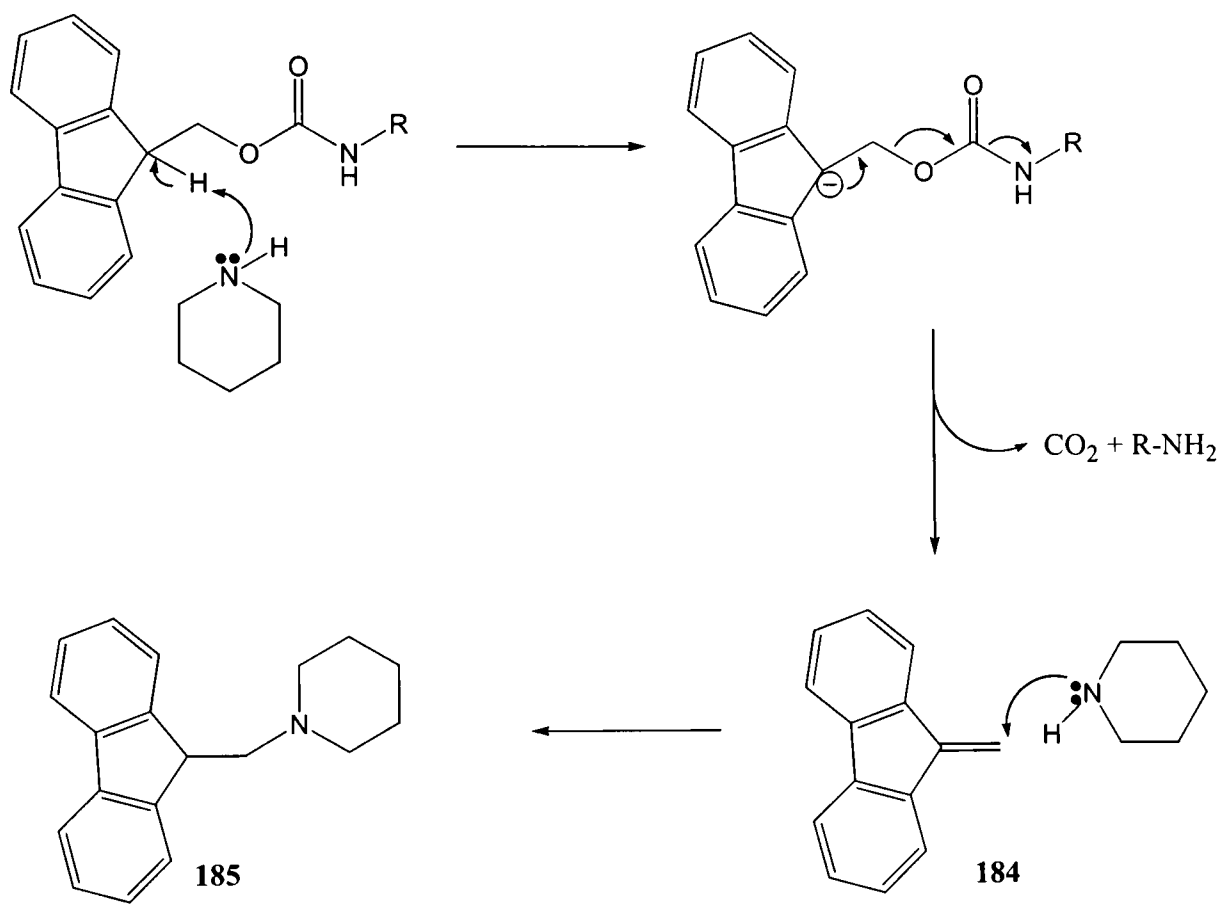
Scheme 52

Fmoc-CA **182** was attached to 2-chlorotrityl resin **181** (substitution 1.08mmol/g) under standard conditions[143]. A solution of diisopropylethylamine (DIPEA) and **182** in dry DCM was added to the resin under nitrogen. Filtration and washing with solvent afforded the resin bound Fmoc-CA. After this, and every subsequent coupling step, a test of Fmoc-substitution was performed to see if the coupling reaction was complete, i.e. that there were no unreacted groups which could react in future coupling reactions to form truncated peptides. The theoretical substitution can be calculated using Equation 10[145].

$$A = \frac{B \times 1000}{[1000 + (B \times (M - X))]}$$ **Equation 10**

Where A is the theoretical substitution in mmol/g, B is the substitution of the starting resin, M is the molecular weight of the peptide and protecting groups, and X has a value of 36 for a chlorotrityl resin. The actual substitution can be measured by reacting the Fmoc protected resin with a solution of 20% piperidine in DMF. Piperidine cleaves the Fmoc protecting group *via* a β-elimination process to form a dibenzofulvene product

184 (Scheme 53). Excess piperidine adds to this product to form a fulvene-piperidine by-product **185** which has a characteristic UV absorbance at 290nm. If a known amount of the substituted resin (~1mg) is reacted with a known volume of 20% piperidine in DMF (3ml), and the absorbance at 290nm measured, then using Equation 11[145], which has been derived from the Beer-Lambert Law (ε for the piperidine fulvene adduct is $5252 M^{-1}cm^{-1}$), the substitution of the resin in mmol/g can be calculated.



Scheme 53

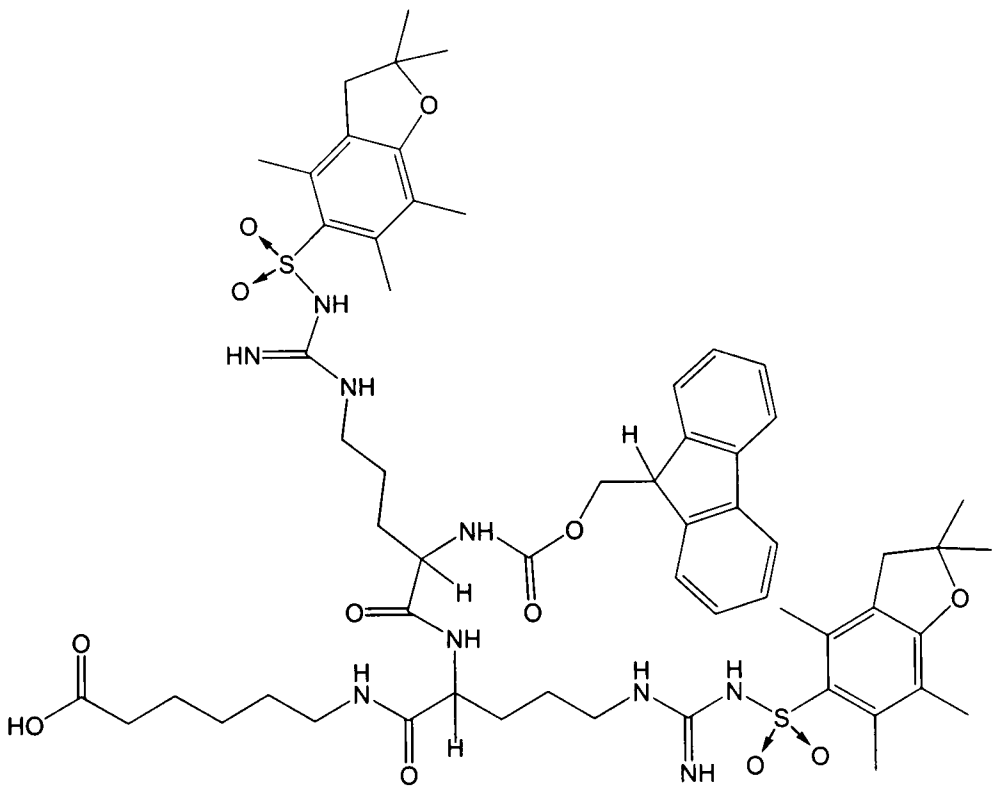
$$\text{Substitution} = \left(\frac{\text{absorbance}}{\text{mass of sample x } 1.75} \right) \qquad \textbf{Equation 11}$$

The results of the tests for this, and the subsequent coupling steps are summarised in Table 7.

Coupling Step	Theoretical Substitution	Result of Fmoc Analysis
Attachment of FmocCA to resin	0.81mmol/g	0.88mol/g
1st Arginine coupling	0.59mmol/g	0.6mmol/g
2nd Arginine coupling	0.38mmol/g	0.3mmol/g

Table 7

The Fmoc protecting group was removed by sonicating the resin in 20% piperidine in DMF. The first FmocArg(Pbf) residue was coupled to the resin by mixing with DIPEA, PyBOP and HOBt in DMF. Fmoc deprotection as before, followed by the coupling of the second FmocArg(Pbf) residue afforded the protected peptide attached to the 2-chlorotrityl resin. The peptide was cleaved from the resin using a very mild cleavage cocktail of 2:2:6 acetic acid:trifluoroethanol:DCM. Chromatography yielded the protected peptide **186** (Figure 36) ready for incorporation into a polymer.



186

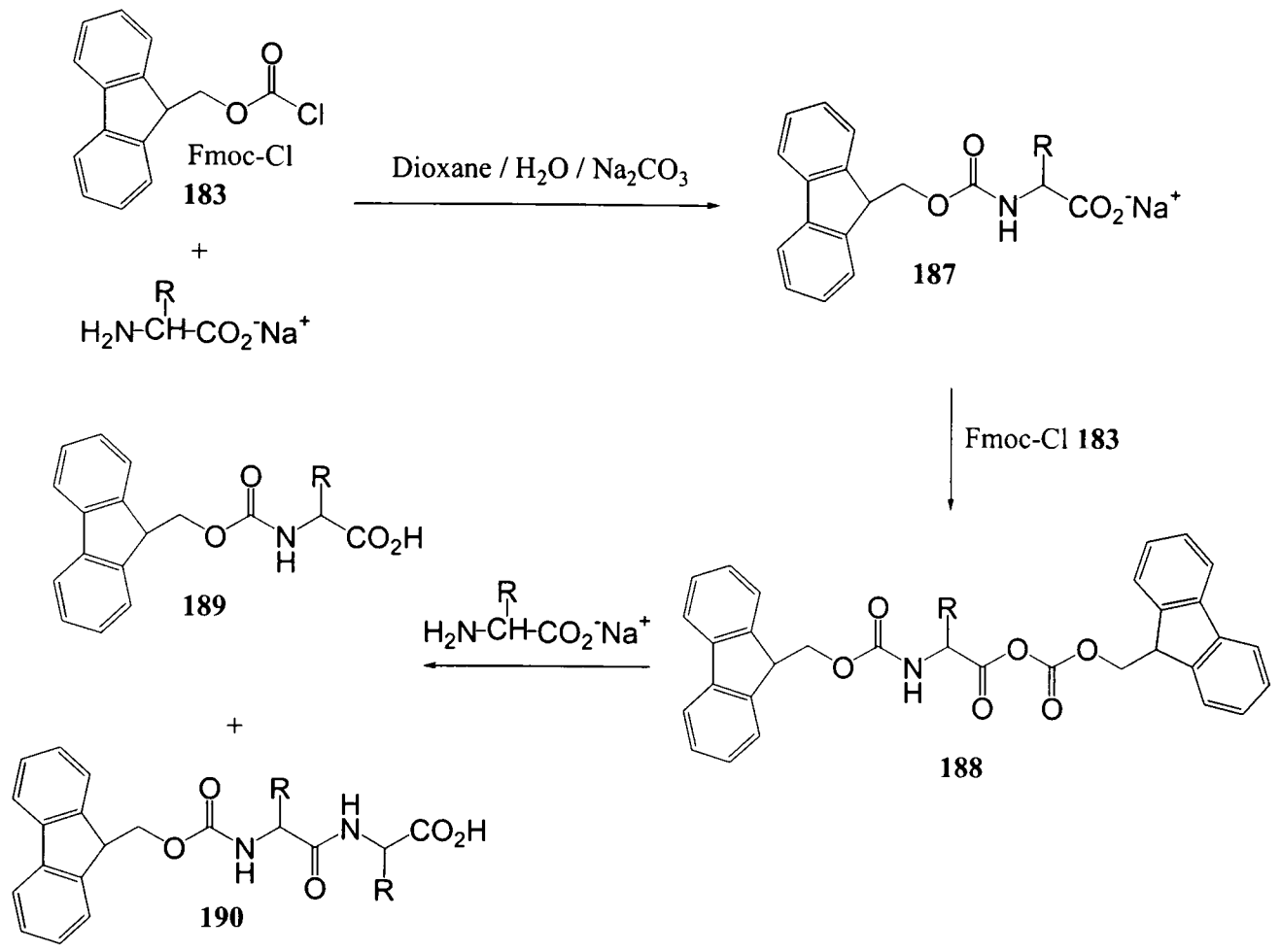
Figure 36

7. Synthesis of the Catalytic Polymers.

With the binding "receptor unit" in hand, we were now ready to synthesise the catalytic polymers. Polyallylamine hydrochloride is commercially available in two molecular weights- ~11,000kDa, and ~70,000kDa. For our reactions we decided to utilise the higher molecular weight polyallylamine. The molecular weight of the monomer in the hydrochloride salt is 93, so each high molecular weight polymer contains approximately 752 amino groups. When calculating quantities and equivalents of reagents, the number of monomer equivalents in the polymer was always used. As discussed in section 5, any unreacted amino groups on the polyallylamine backbone were to be functionalised using Fmoc protected lysine. Since this meant that one third of all the amino groups on the polymers would contain lysine, a large quantity of the protected lysine starting material would be required, and although this is commercially available, it was more economical to synthesise it in house.

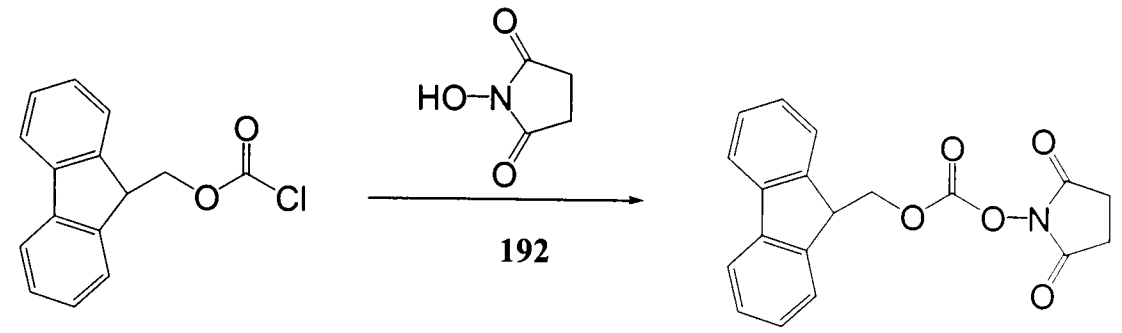
7.0 Synthesis of Fmoc protected lysine.

Traditionally, Fmoc protecting groups are introduced by reacting the amino acid with Fmoc-Cl **183** under basic conditions[144]. Under these conditions however, formation of a dipeptide by-product has been seen to occur (Scheme 54)[146]. The Fmoc-amino acid anion **187** formed as the initial product is able to act as a nucleophile and attack the Fmoc-Cl **183** to form a mixed anhydride **188**. The amino acid starting material can then either attack the carbonyl group of the O-Fmoc moiety to afford the desired product **189**, or the carbonyl of the amino acid group, to afford an Fmoc-dipeptide by-product **190**.

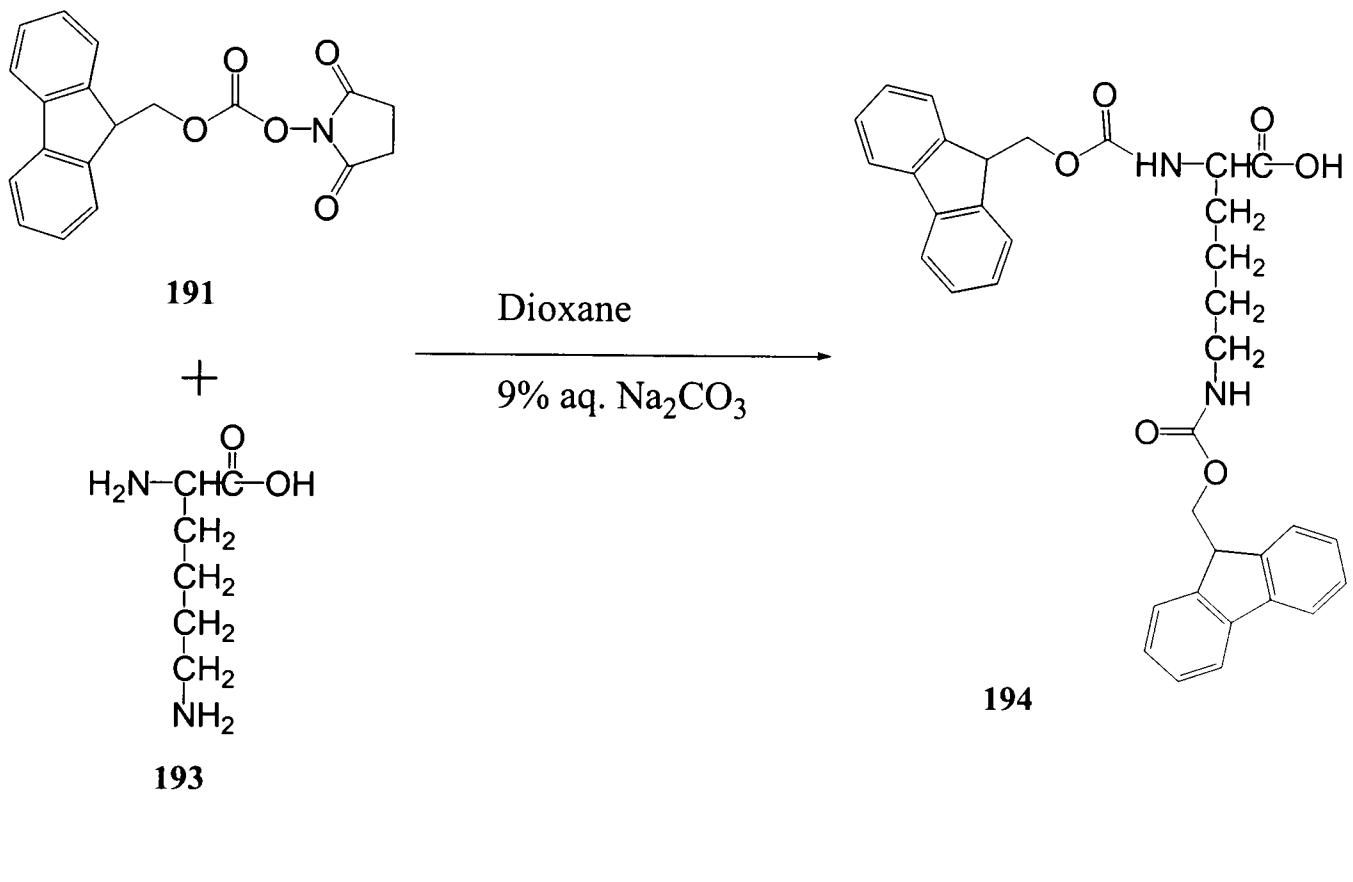


Scheme 54

The use of succinimidyl 9-fluorenylmethyl carbonate (Fmoc-OSu) **191**, which is readily prepared from Fmoc-Cl **183** (Scheme 55), has been reported to avoid the formation of this by-product[146]. This reagent seemed a better choice for the synthesis of protected lysine. Fmoc-OSu **191**was prepared in a 90% yield using a literature procedure by reacting Fmoc-Cl **183** with *N*-hydroxysuccinimide (NHS) **192** and triethylamine in dioxane. Addition of Fmoc-OSu **191** to a solution of lysine **193** in 9% aqueous sodium carbonate solution, afforded the Fmoc protected lysine. Purification by washing with water yielded pure **194** in 71% yield (Scheme 55).



183 **191**

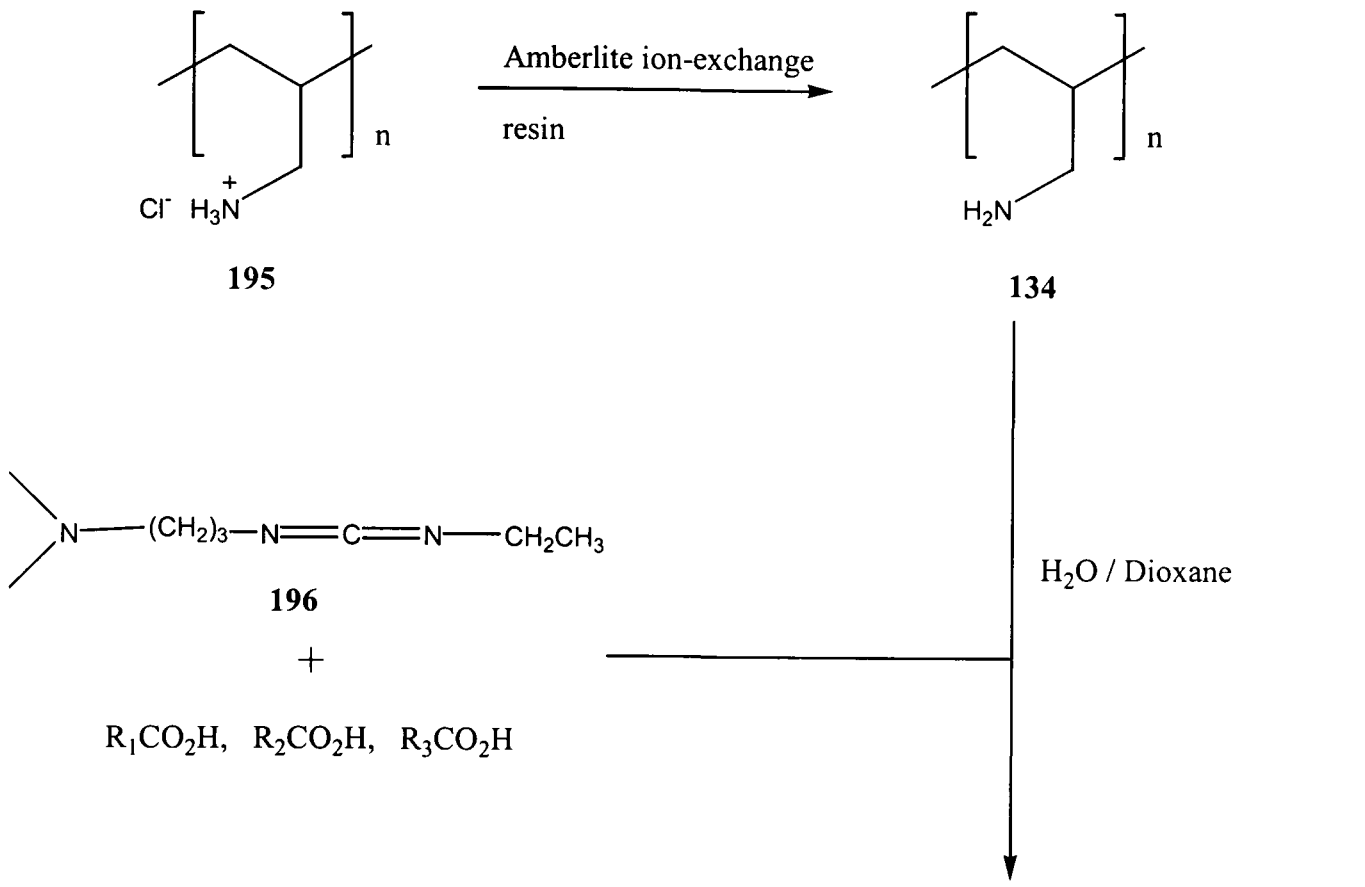


193

Scheme 55

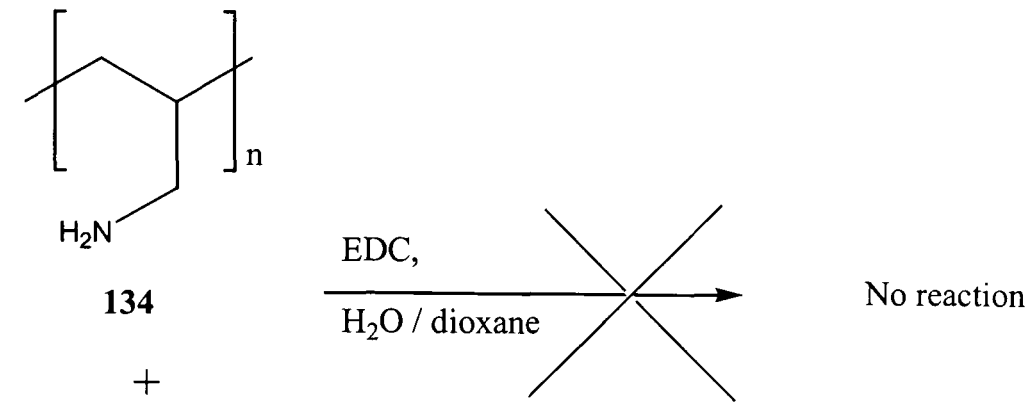
7.1 Synthesis of Test and Control Polymers.

For the attachment of the receptor site and functional groups to the polymer backbone we had envisaged using the methodology described by Menger in his synthesis of artificial phosphatases (Scheme 56)[95]. Polyallylamine hydrochloride **196** was converted to the free base **134** by stirring with an ion-exchange resin. The free amine **134** was then reacted with a range of carboxylic acids and 1-[3-(dimethylamino)propyl]-2-ethylcarbodiimide hydrochloride (EDC) **196** in either dioxane or water to afford the functionalised polymer **197**.



Scheme 56

The simplest polymer to be synthesised was a control containing only lysine and the polyallylamine backbone. It was logical therefore, to attempt the synthesis of this polymer first. Following Menger's approach, when 1.5 equivalents of FmocLys(Fmoc)OH **194** and EDC **196** were heated with polyallylamine **134** at 50°C for 2 days, no coupling between the polymer and amino acid was observed (Scheme 57). The bulky aromatic Fmoc groups make **194** relatively non-polar, and consequently insoluble in water and dioxane. We therefore reasoned that a change of reaction solvent was necessary for the reaction to proceed.



134

+

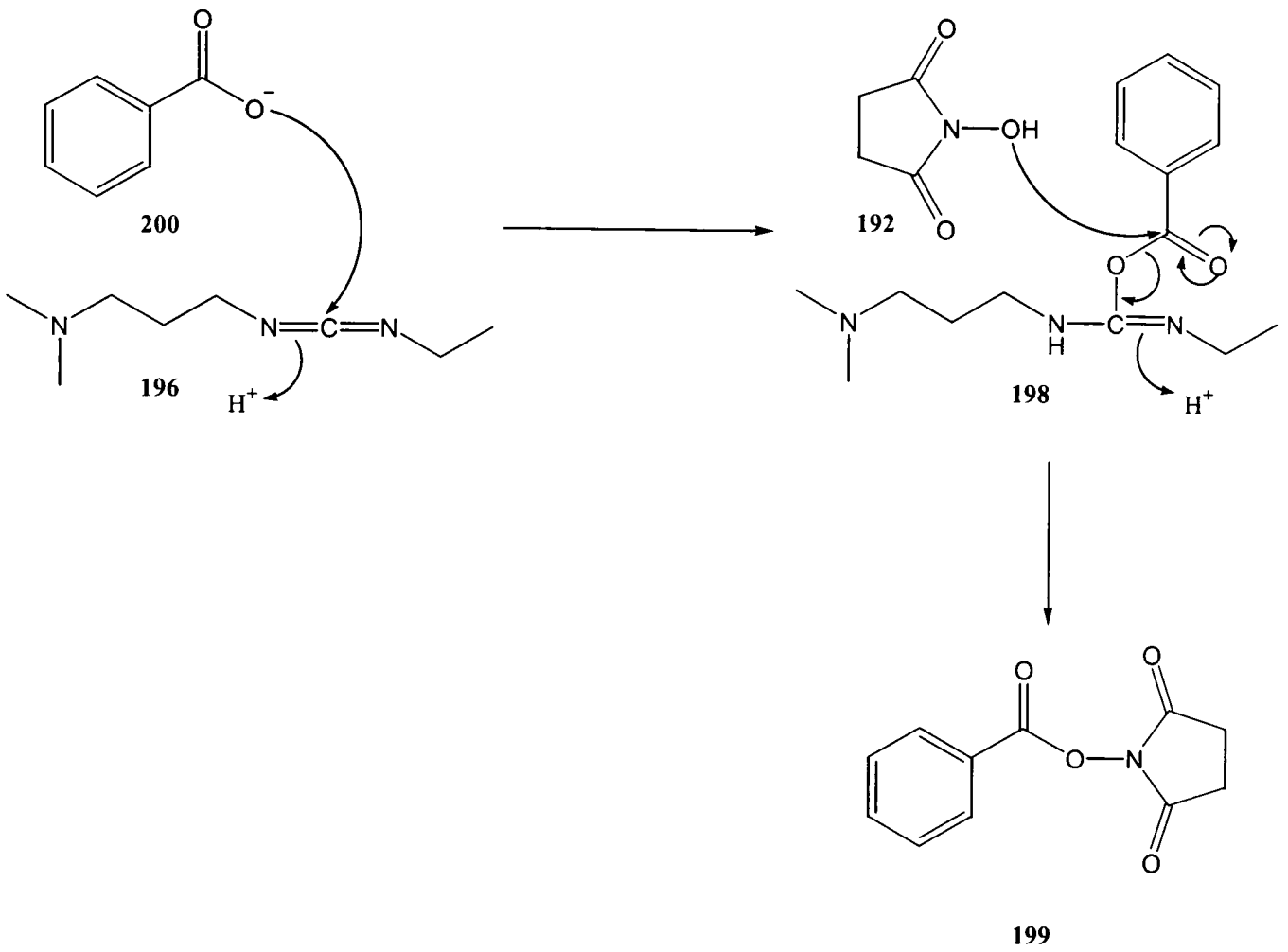
FmocLys(Fmoc)OH

194

Scheme 57

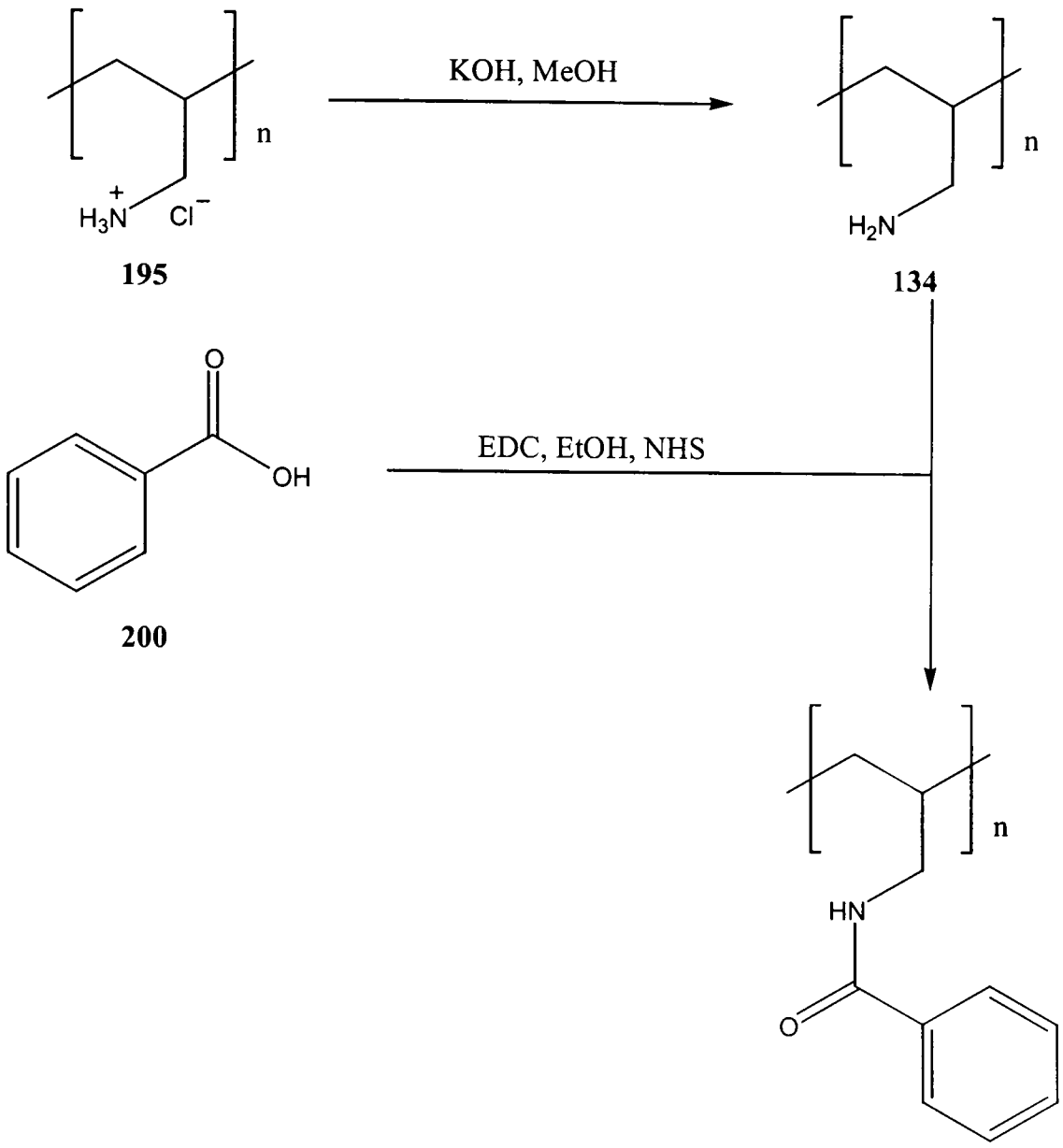
Tanaka *et al.* have recently published work in which they use polylallylamine derivatives as stationary phases for electrokinetic chromatography[147]. This was of interest as they used a different method to convert the polyallylamine hydrochloride to the free base. They report that if polyallylamine hydrochloride **195** is stirred in methanol with 1 equivalent of potassium hydroxide overnight, the free base is formed. Removal of solvent followed by the polymer being redissolved in ethanol, precipitates the potassium chloride formed during the course of the reaction and leaves the free base **134** ready to use in an organic solvent. They have then performed coupling reactions onto the polymer using a variety of different solvents.

We decided to use this methodology for our reactions since the use of organic solvents was desirable. It was also decided to add NHS **192** to the reaction mixture as an additional activating agent[148]. Activation of the carboxylic acid is by attack of the acid on the carbodiimide moiety in **196**. The hydroxyl group of NHS can then attack the intermediate **198** to form an active ester **199** which was hoped would be more susceptible to nucleophilic attack from the polyallylamine than the product **198** formed from the reaction between benzoic acid and EDC (Scheme 58).



Scheme 58

Initially this method was tested using a system comprised of polyallylamine **134** and benzoic acid **200** in ethanol (Scheme 59). A solution of polyallylamine **134** in ethanol was prepared as described above. Benzoic acid **200** was dissolved in ethanol, and 1.2 equivalents of EDC **196**, and 1.4 equivalents of NHS **192** were added, and the solution stirred for 4 hours. The reaction was monitored using TLC so that when no further reaction between the benzoic acid **200**, EDC **196** and NHS **192** was observed, a solution of polyallylamine **134** in ethanol was added to the mixture. A white precipitate instantly formed, which was collected by filtration, to afford a white plastic-like gum. As the gum would not dissolve in any solvent, it was analysed using solid state NMR which showed that the product was **201** with the benzoic acid coupled to the polymer as we had hoped.

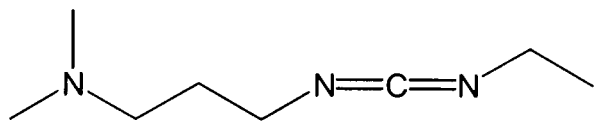


195 **134** **200** **201**

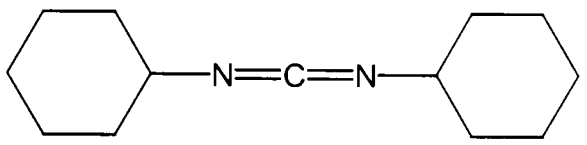
Scheme 59

Although the test reaction was successful, a couple of changes had to be made to the reaction conditions before they could be applied to the reaction between polyallylamine **134** and FmocLys(Fmoc)OH **194**. Firstly, **194** proved to be insoluble in ethanol, and further investigations found solubility to be highest in chloroform. No previous syntheses of polyallylamine derivatives have reported the use of chloroform as a solvent. It was found however, that free polyallylamine could be completely solvated by chloroform, which was therefore used as the solvent for the reaction between the amino acid residue and activating agents. The free polyallylamine was still added to the reaction as an ethanolic solution.

Secondly, the use of chloroform as a solvent made EDC a poor choice of activating agent. EDC **196** was developed for use in coupling reactions in which the reaction proceeds in aqueous media[149]. For use in organic solvents such as chloroform, the polar nature of EDC makes it unsuitable, and the less polar *N, N'*-dicyclohexylcarbodiimide (DCC) **202** (Figure 37) is a more suitable choice of reagent, as not only is it completely soluble, it is also more economical[150].



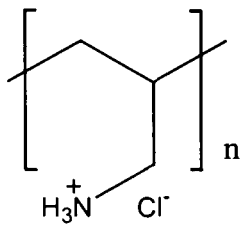
196



202

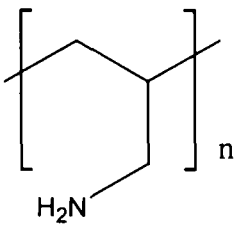
Figure 37

A solution of polyallylamine free base **134** in ethanol was prepared as described above and added to a solution containing 1 equivalent of FmocLys(Fmoc)-OH **194**, 1.2 equivalents of DCC **202** and 1.4 equivalents of NHS **192** in chloroform (Scheme 60). After stirring for 3 hours, a white solid had precipitated out of solution. As with the polyallylamine derivative formed from benzoic acid, the white solid proved to be insoluble in any solvent, and analysis was performed using solid state ^{13}C NMR, and solid state IR. If the coupling had occurred, the NMR data should show two different types of carbonyl peaks relating to the amide bond where the amino acid is coupled to the polymer, and the carbamate linker of the Fmoc group. The NMR data indeed showed an amide peak at 174.5ppm, and a peak at 157.6ppm, which is characteristic of the carbonyl moiety in the Fmoc group. IR analysis showed the presence of two different carbonyl groups, at 1693 and 1651cm^{-1}. The two analyses combined clearly indicated that the coupling between the amino acid residue and polyallylamine to form polymer **203** had proceeded cleanly and efficiently.



195

KOH, MeOH



134

EtOH

FmocLys(Fmoc)-OH **194**

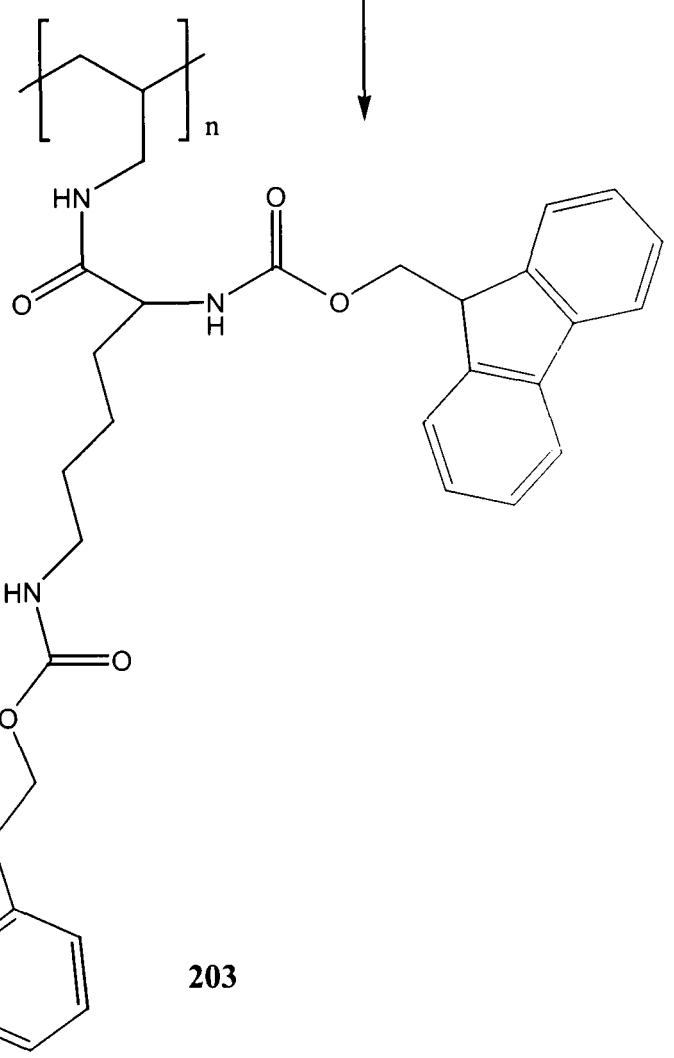
\+

DCC **202**

\+

NHS **192**

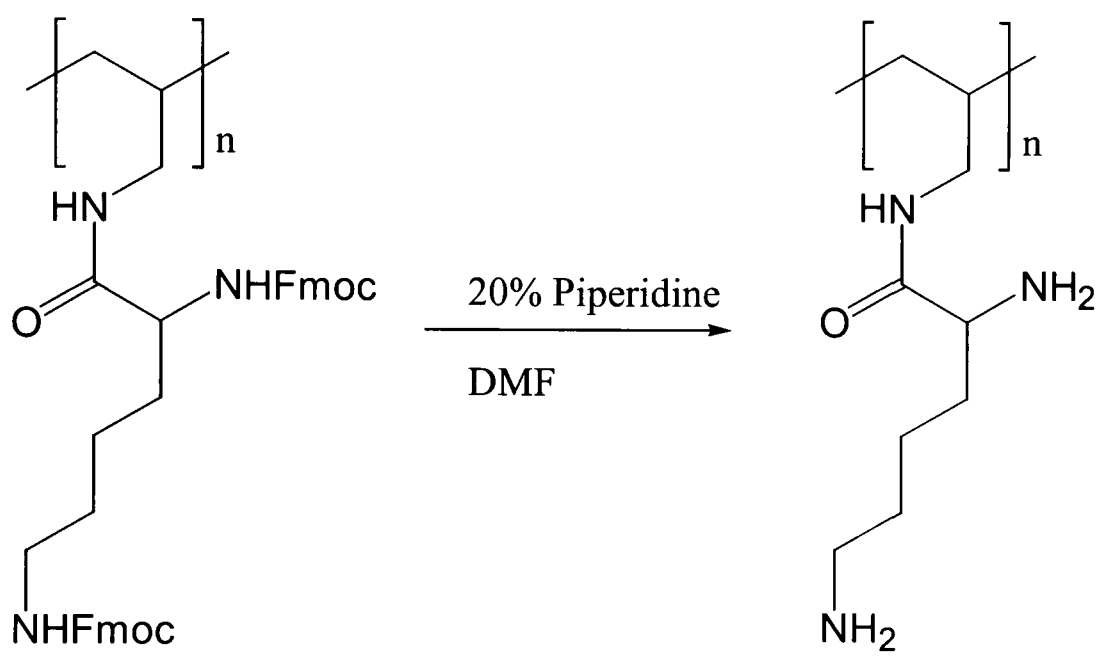
$CHCl_3$



203

Scheme 60

Deprotection of polymer **203** proved to be straightforward. The polymer was suspended in a solution of 20% piperidine in DMF, and sonicated for 3 hours. Filtration, and extensive washing with DMF and DCM, followed by drying afforded the deprotected polymer **204** as a white solid (Scheme 61). When the solid was added to acidic D_2O, a transparent gel formed, which could be analysed by solution phase NMR. ^1H NMR showed an aliphatic polymer, with no peaks in the aromatic region of the spectra, thus showing that the Fmoc groups had been cleanly removed.

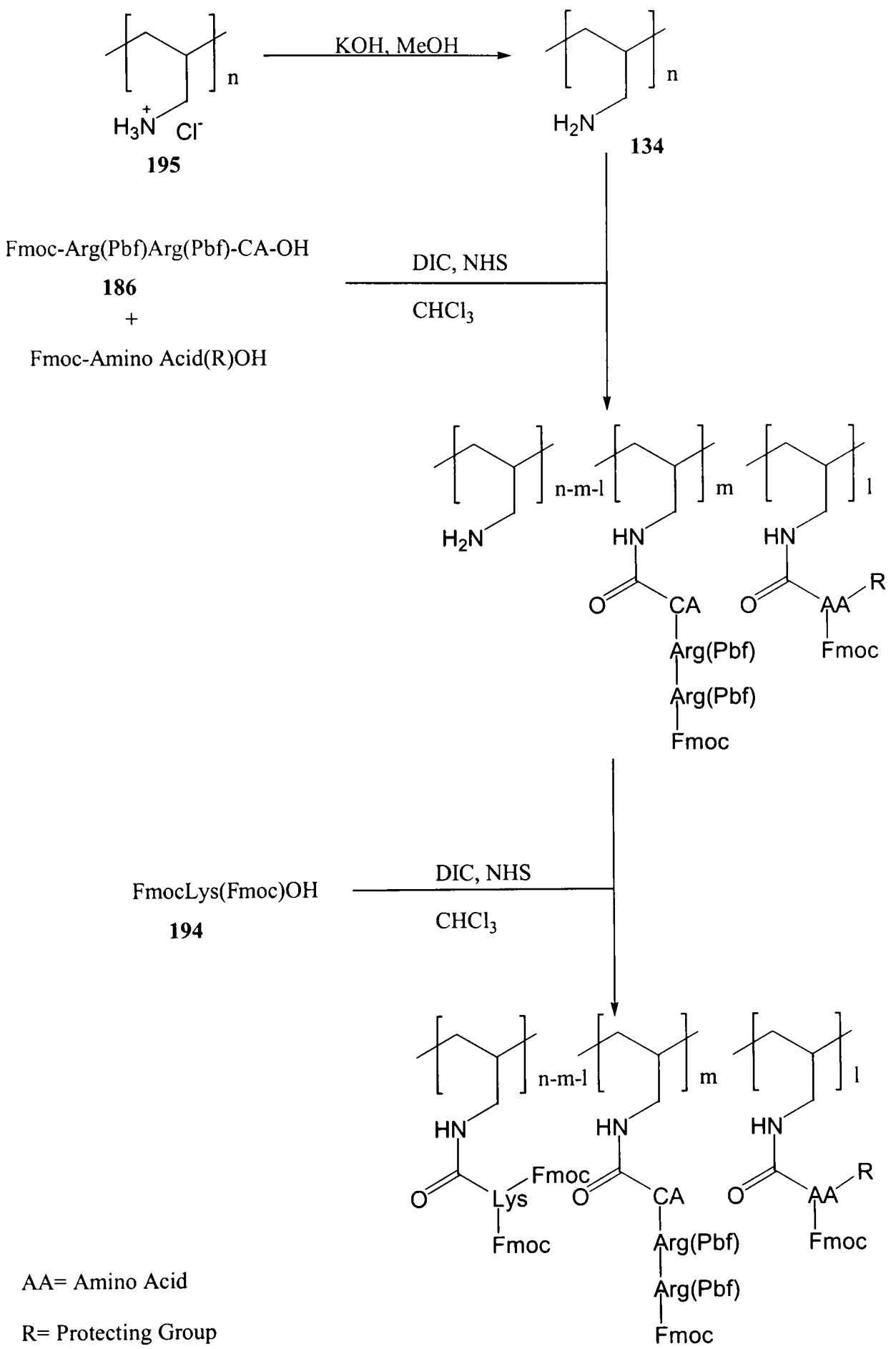


203 **204**

Scheme 61

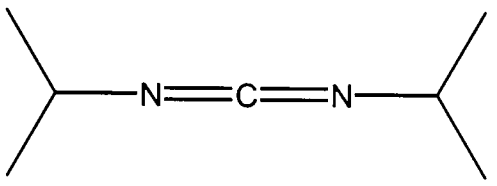
7.2 Synthesis of the Functionalised Polymers.

With the control polymer successfully synthesised it was time to apply the methodology to the more complex polymers containing more than one functional group. A general synthesis was designed as shown in Scheme 62. It was vital that both the receptor unit containing the Arg-Arg moiety **186**, and the functional amino acid residues were incorporated into the polymers before the remaining free amino groups on the polymer backbone could be functionalised with lysine. Therefore they would be coupled first onto the polyallylamine backbone. Once TLC has showed this reaction to be complete, the lysine could then be coupled to the polymer. The Arg-Arg peptide **186** and amino acid "catalytic group" being coupled simultaneously should influence each other's position on the polymer backbone.


KOH, MeOH
n
H3N+ Cl-
195
H2N
134
Fmoc-Arg(Pbf)Arg(Pbf)-CA-OH
186
+
Fmoc-Amino Acid(R)OH
DIC, NHS
CHCl3
n-m-l
m
l
HN
O
CA
AA
R
Arg(Pbf)
Arg(Pbf)
Fmoc
FmocLys(Fmoc)OH
194
Lys
AA= Amino Acid
R= Protecting Group

Scheme 62

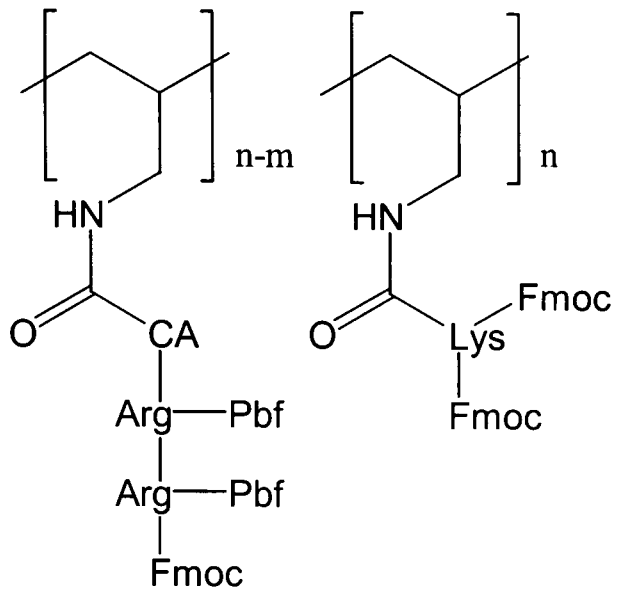
The activating reagent was again altered, this time from DCC **202** to *N,N'*-diisopropylcarbodiimide (DIC) **205**. Ureas are formed as a by-product of the reaction when carbodiimides are utilised as activating agents. When DCC is used the urea formed is highly insoluble in organic solvents, and thus is hard to remove from the reaction product. DIC forms a soluble urea which can be removed by washing, and hence was the preferred reagent[151].



205

Figure 38

The first polymer synthesised in this manner was **206** (Figure 39), and contained only the Arg-Arg peptide and lysine. A solution containing 0.3 equivalents of FmocArg(Pbf)-Arg(Pbf)OH **186** and DIC **205** and NHS **192** in chloroform was stirred and monitored by TLC. A solution of polyallylamine free base **134** in ethanol was then added to the mixture. No white precipitate formed in this instance. The Pbf group stains a distinct pink colour with anisaldehyde. When TLC was run, a new spot which stained this distinct pink colour had appeared on the baseline, and failed to move even when the eluent was changed to 100% methanol. This suggested that a high molecular weight unit had been formed which contained the Pbf group. When TLC showed no further reaction, a solution containing 0.6 equivalents of FmocLys(Fmoc)OH **194**, DIC **205** and NHS **192** in chloroform was added to the reaction, and the resulting mixture left to stir for 48 hours.



206

Figure 39

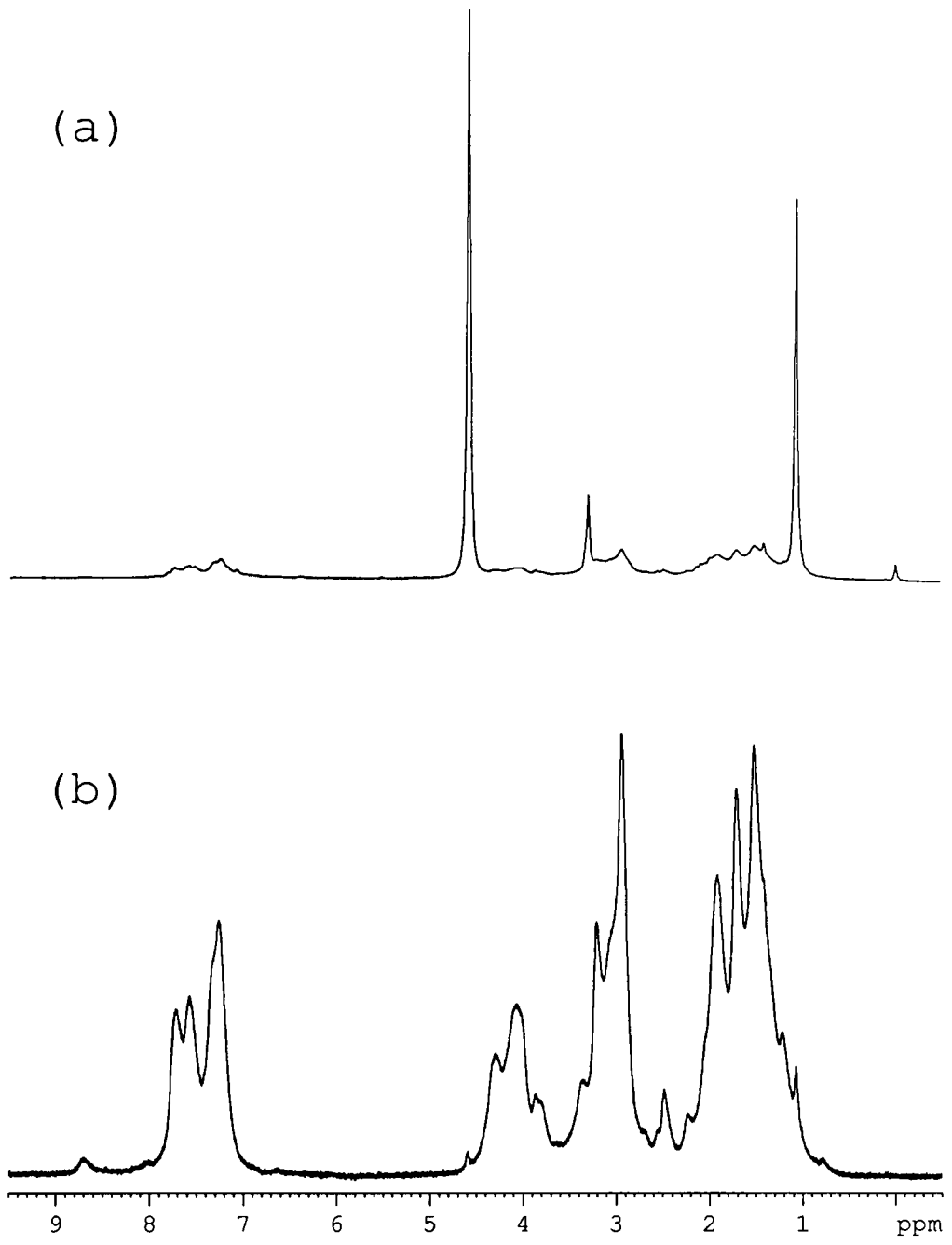
Disappointingly after this time, only a small amount of white precipitate had formed, which was collected by filtration. NMR analysis showed it to be a low molecular weight material. Attention therefore turned to the chloroform solution. TLC still indicated the presence of the coupled polymer product in the solution. Dilution of this solution with methanol resulted in the immediate precipitation of a white gum, which was collected by filtration and washed repeatedly with methanol. The solubility of the polymer in chloroform meant that the product could now be analysed using solution NMR.

7.3 Polymer Analysis using PFG NMR.

When the solution ^{1}H NMR of the polymer was first run, only solvent peaks appeared to be present in the spectrum (see Figure 40 a). Upon closer inspection however, it was evident that there were other very broad peaks present which were masked by the sharper solvent peaks. NMR spectra of high molecular weight molecules appear as broad lines due to efficient relaxation[152].

The masking of the broad polymer peaks by the sharper peaks of impurities could prove problematic. An initial glance at the spectrum would indicate that there is a high percentage of impurities present in the sample due to the high intensity of the sharp peaks. In fact, this is more likely to be a simple result of the much higher solubility of low molecular weight species. In order to view true spectra of our polymers, we needed an NMR technique that would effectively mask the peaks of the low molecular weight impurities. We decided to continue our use of PFG NMR techniques to achieve this.

As described on page 95, translational diffusion rates in solution for high and low molecular weight species are significantly different. The intensity of a signal observed in PFG-NMR experiments is dependent on how far the molecule has diffused between the two pulses. Clearly, a low molecular weight species will have travelled further. Hence at high gradient strength ($25Gcm^{-1}$), signals corresponding to low molecular weight species will be suppressed, and the signals of the polymer can be observed more clearly. This technique was used on all of the polymer samples. An example of the effectiveness of the method can be seen in Figure 40.

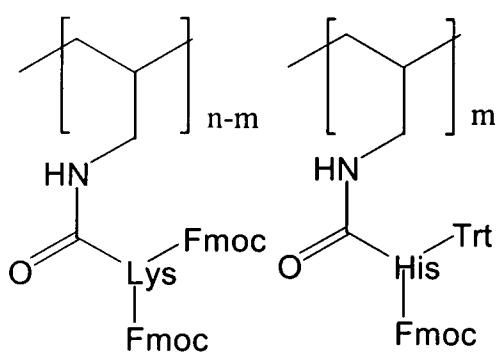


a) Polymer **206**. NMR spectra no gradient applied. b) NMR spectra with 50% gradient.

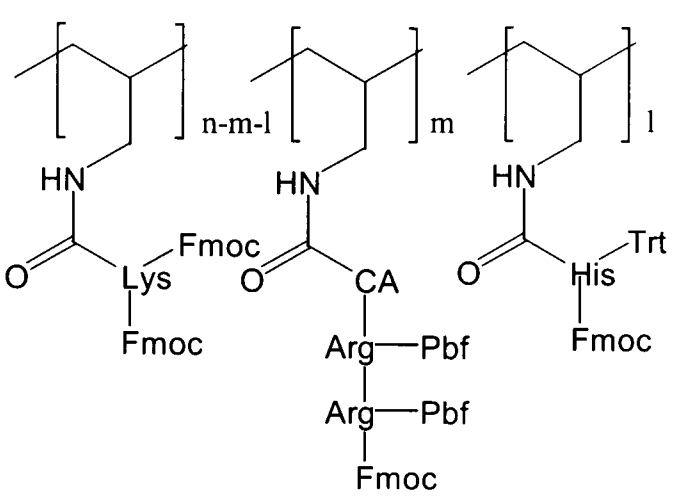
Figure 40

7.4 Synthesis of Other Polymers.

Further polymers were synthesised using the methodology described above. Initial attention was focused on the synthesis of polymers containing histidine residues as the catalytic group. Two polymers were successfully synthesised, one, **207** containing only histidine and lysine in a 1:2 ratio, and the other **208** containing lysine, histidine and the Arg-Arg receptor unit in 1:1:1 ratios, which was of course the first example of the ultimate objective of our modular strategy (Figure 41).



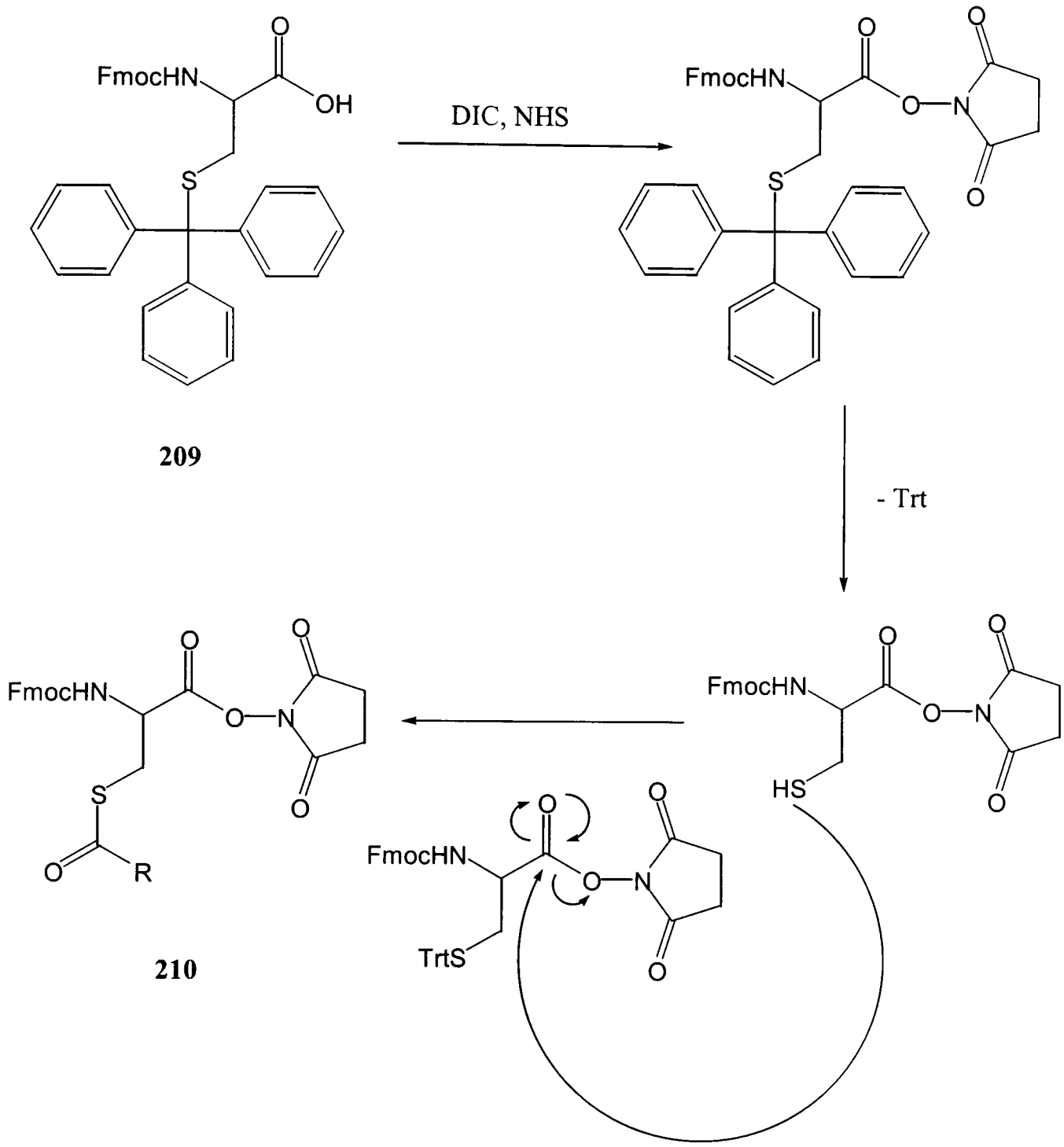
207



208

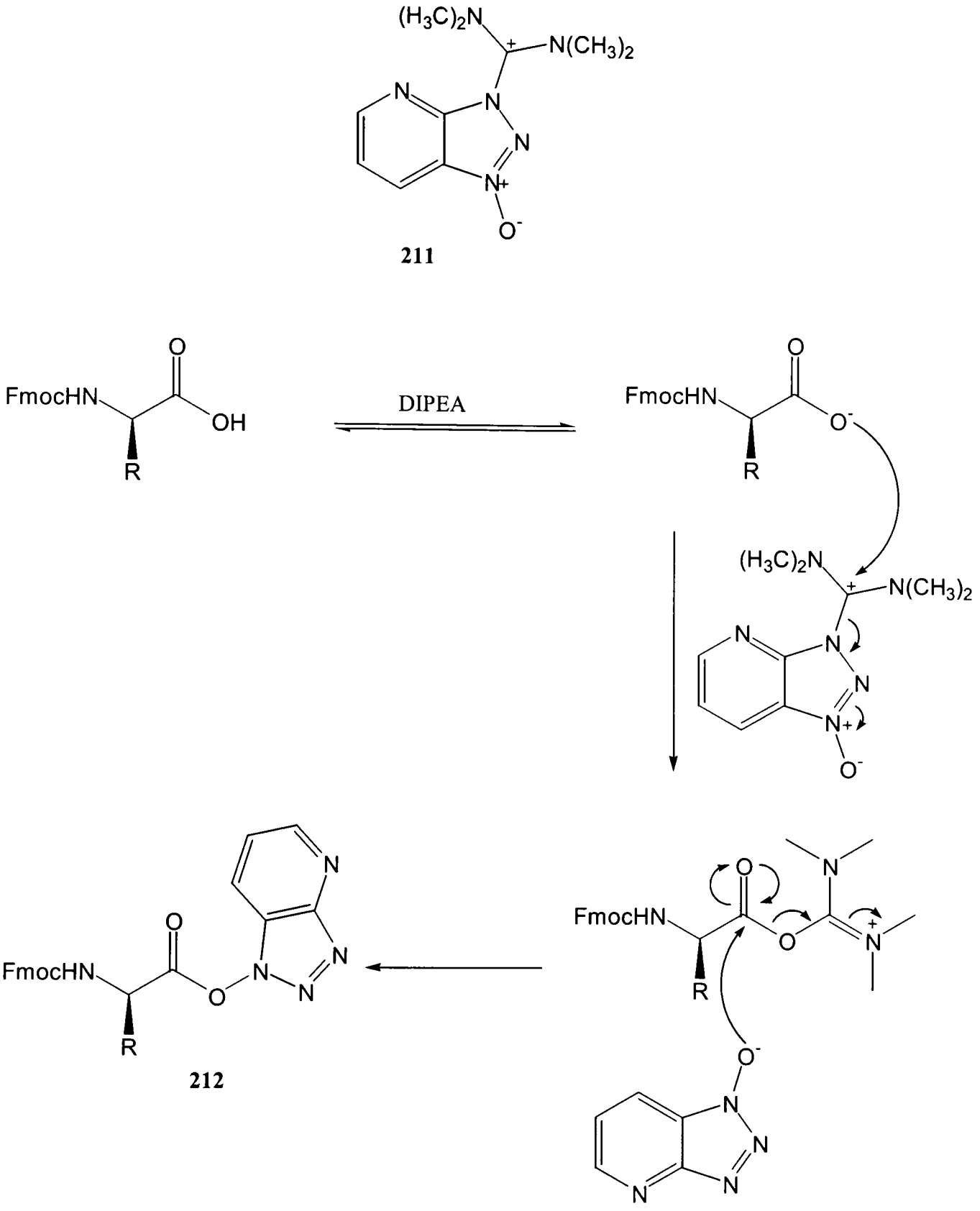
Figure 41

Polymers containing cysteine residues were the next target. However, when 0.3 equivalents of FmocCys(Trt)OH **209**, and 0.6 equivalents of FmocLys(Fmoc)OH **194** were reacted with DIC **205** and NHS **192**, and added to the polyallylamine solution, none of the desired polymeric product was observed by ^{1}H NMR. Lysine had attached to the polymer as before, but none of the cysteine residue could be observed. The use of the trityl protecting group with cysteine has two problems. Firstly it is easily cleaved, and can fall off during the course of the reaction. In this instance the thiol group is now able to act as a nucleophile, and can attack the activated ester, resulting in the formation of thiol esters such as **210** (Scheme 63). Second, the steric bulk of the S-Trt group can prevent the polymer attacking the activated ester. Racemisation is also a problem in couplings involving cysteine residues[145].



Scheme 63

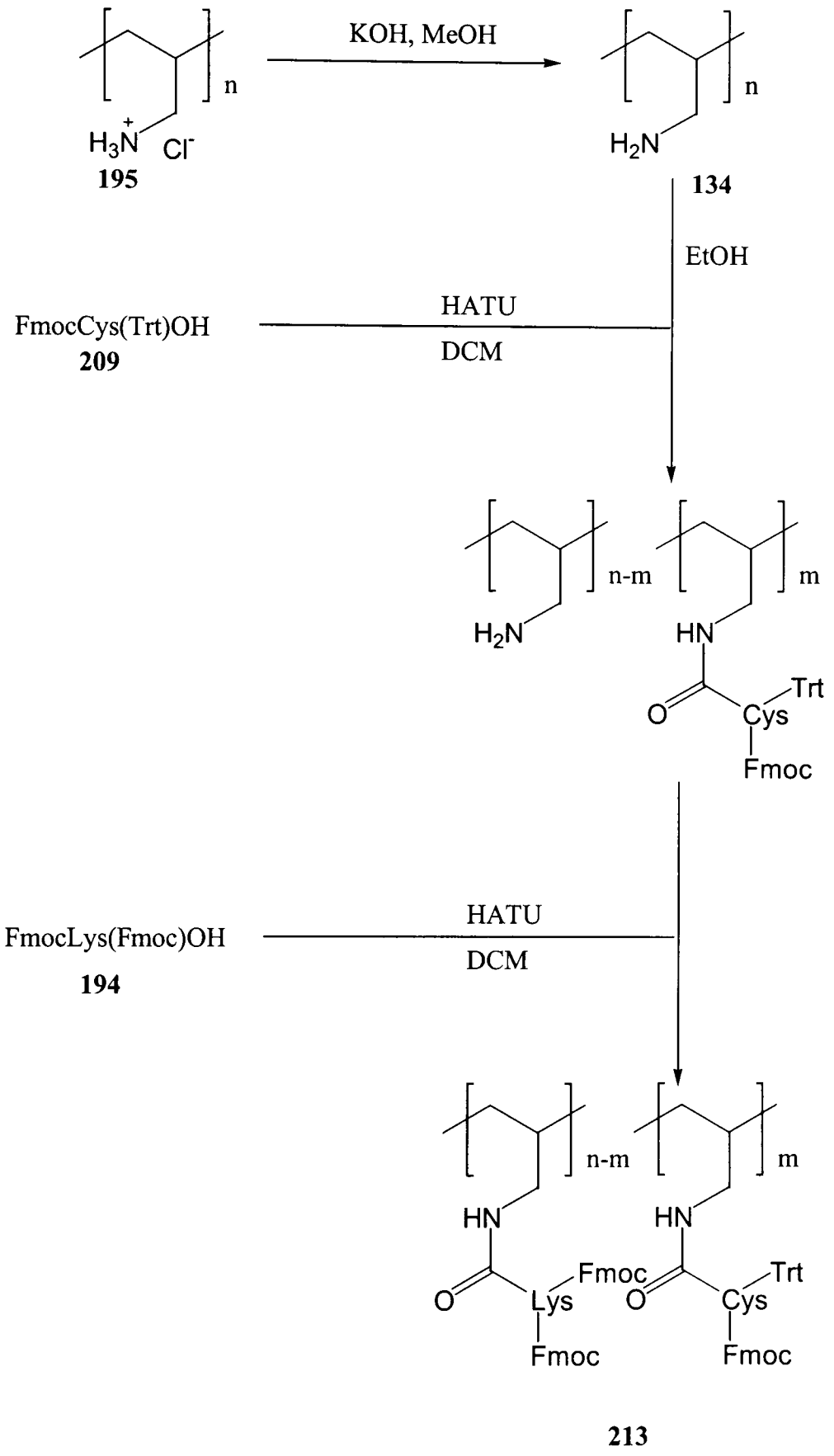
Instead of the DIC/NHS combination, we decided to use a more reactive activating agent. HATU **211**, is an aminium based derivative which when used in the presence of a tertiary base, has been shown to give the most superior results for peptide coupling, as well as suppressing enantiomerization[153]. The mechanism of activation to form the OAt ester **212** is shown in Scheme 64.



Scheme 64

FmocCys(Trt)OH **209** was stirred with 1 equivalent of HATU **211** and DIPEA in DCM and added to an ethanolic solution of polyallylamine **134**. After stirring for 30 minutes, a solution of Fmoc(Lys)FmocOH **194** and HATU **211** in DCM was added to the reaction mixture, which was stirred for a further two days (Scheme 65). In this instance

a white precipitate formed, which was collected and washed, and shown by NMR analysis to be the correct product **213**.



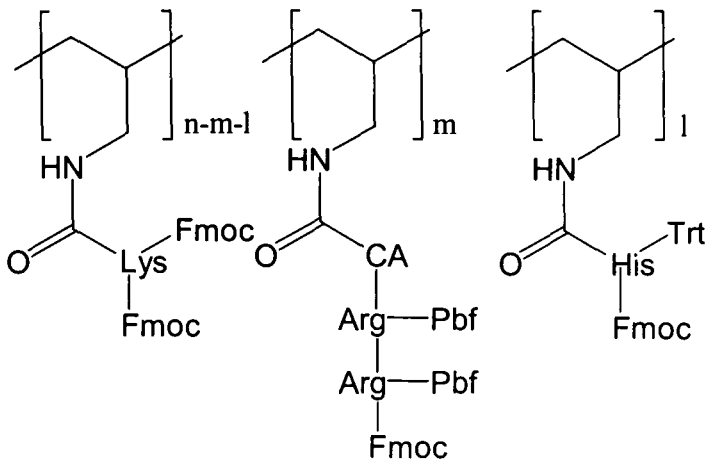
Scheme 65

Unfortunately due to time limitations, we were unable to synthesise any more polymers. However, the use of HATU as the activating reagent shows promise for hindered

systems. For the synthesis of polymers which do not pose problems of steric hindrance of racemisation, the DIC/NHS system would still be more appropriate in view of the expense of the HATU reagent.

7.5 Deprotection of Polymers.

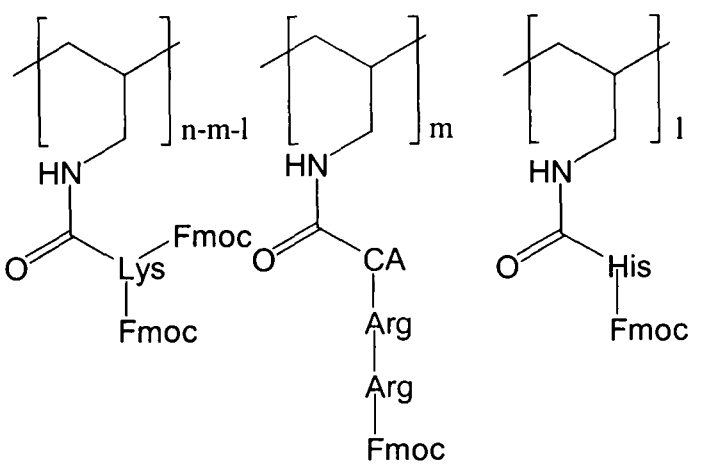
All the polymers were deprotected using the same method demonstrated for polymer **208** in Scheme 66. The polymers contain both acid labile (Pbf, Trt) protecting groups, and base labile (Fmoc) protecting groups. It was desirable for the fully deprotected polymer to be in a neutral form, without any counter-ions present, so it was decided to cleave the acid labile groups first. This was achieved by sonication of the polymer with a mixture of 95:2.5:2.5 trifluoroacetic acid: triisopropylsilane (TIS):H_2O. The role of the TIS was two fold. First it is known to accelerate the cleavage of the arginine Pbf protecting group[154]. Secondly, it acts as an effective quencher for the stable cation formed during cleavage of trityl protecting groups[155]. Filtration, followed by extensive washing with water and DCM afforded the polymers free of acid labile protecting groups.



208

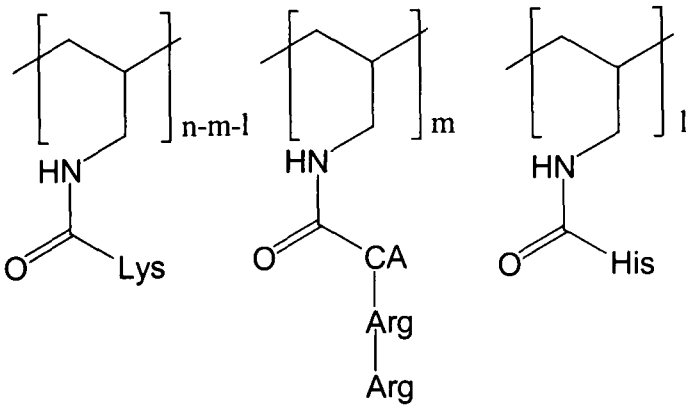
95:2.5:2.5

TFA:TIS:H_2O



214

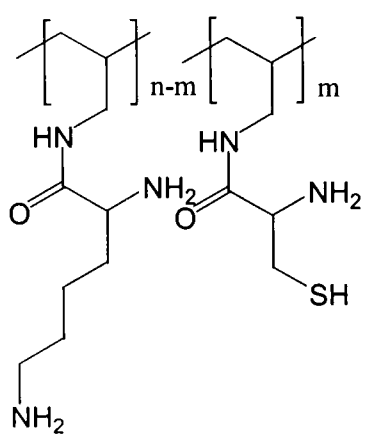
20% Piperidine/DMF



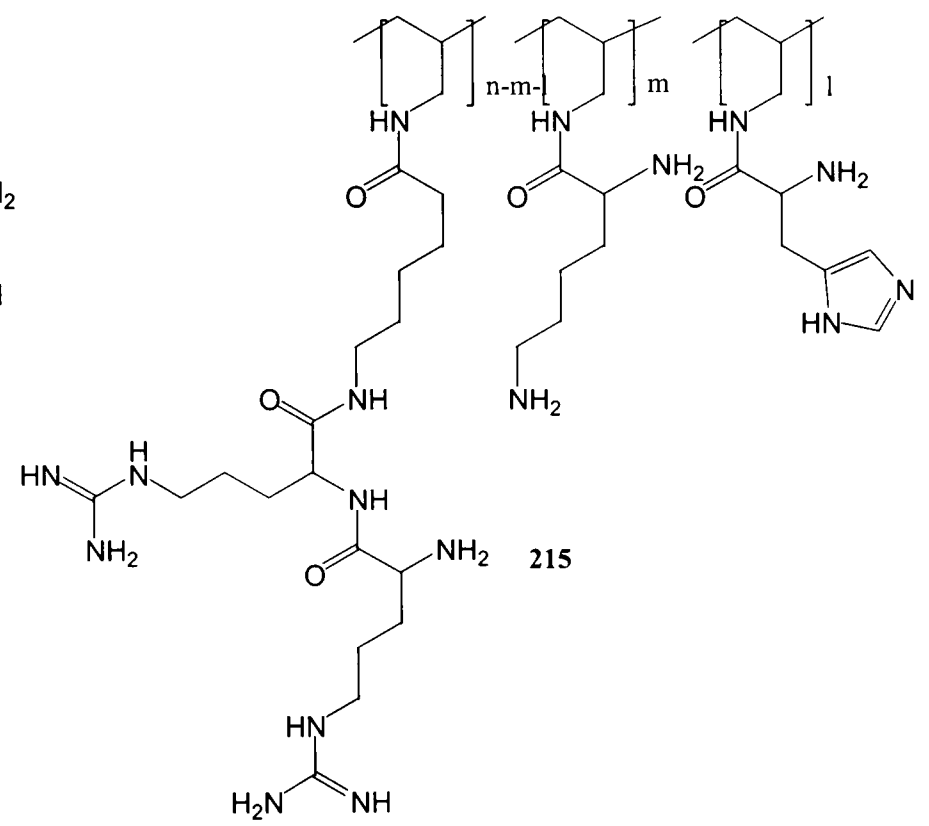
215

Scheme 66

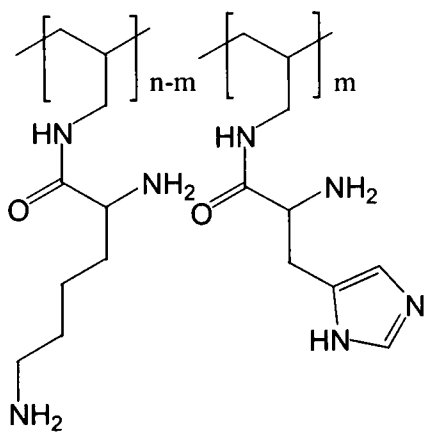
The base labile Fmoc groups were removed by sonication with a 20% piperidine in DMF solution. Filtration, followed by washing with water, and ether, and extensive drying afforded the fully deprotected polymers **215-218** which were now ready to be tested for catalytic activity.



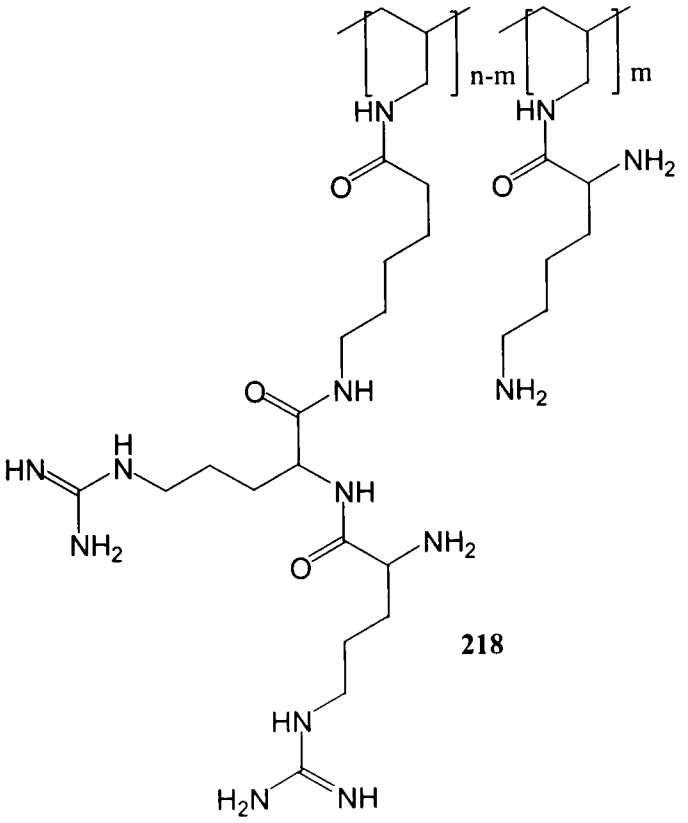
216



215



217



218

Figure 42

8. Binding and Rate Studies of the Polymers.

8.0 Diffusion Rate Measurements of the TSA in the Presence of Polymer.

With the Arg-Arg peptide now incorporated into a polymer, it was important to see if the TSA still showed a strong affinity for the receptor unit now that the surrounding microenvironment had changed. As the role of the Arg-Arg was to stabilise and bind the transition state of the ester hydrolysis, the affinity of the polymer for the TSA is crucial, and if after incorporation into the polymer, the chosen receptor no longer showed any such affinity, then the catalytic capabilities of the polymer would be seriously called into question.

The binding of the TSA **146** to polymer **218**, containing Arg-Arg and lysine, was monitored by measuring the change in diffusion coefficient of the TSA upon complexation. The method used was as described in chapter 4. In this instance however, the difference in molecular weights between the polymer host (MW > 70,000), and TSA ligand (MW 419) is far greater than that between the TSA and the dipeptides used in the earlier studies, and it was expected that the change in diffusion coefficient observed upon binding should be much greater.

Three solutions were prepared. One contained 8.5mg of polymer **218** in 1ml of D_2O, another was a 2mM solution of TSA **146** in D_2O. The third solution consisted of 8.5mg of polymer in 1ml of a 2mM solution of TSA. All solutions were at pD7. The diffusion coefficients of polymer and TSA in all solutions were measured, and the results are shown in Table 8.

Sample	D_{TSA} x $10^{-10}m^2/s$	$D_{polymer}$ x 10^{-10} m^2/s
8.5mg Polymer **218**	–	0.161
2mM TSA **146**	4.251	–
2mM TSA **146** + 8.5mg Polymer	0.861	0.161

Table 8

As can be seen from Table 8, when the TSA was mixed with polymer **218**, the observed diffusion coefficient of the TSA decreased by a value of 3.4 x $10^{-10}m^2/s$. This is a very significant change, and suggests very strong complexation between the TSA and polymer. It was a concern that the change might be a result of the increased viscosity of the solution caused by the presence of the polymer. In order to determine whether or not this was the case, the diffusion coefficient of the water (D_{solv}) was measured. This has a known value of 2.3 x $10^{-9}m^2/s$[156,157], and should the change in diffusion coefficient be due to a change in viscosity of the solution, then the diffusion coefficient of water should also be altered. Fortunately, the measured value of D_{solv} was the same as the literature value, and thus it can be assumed that the observed change in D_{TSA} is due to complexation to the polymer.

As discussed previously, the observed diffusion coefficient, D_{obs}, is actually a mole fraction weighted average of the diffusion coefficient of bound and free molecules (Equation 4). As the molecular weight of the polymer is now considerably larger than the ligand, the assumption can be made that the diffusion coefficient of the complex (D_{COMP}) is the same as that for the free polymer measured and shown in Table 8, and thus the percentage of bound TSA can be calculated from Equation 4.

$$D_{obs} = X_{TSA}D_{TSA} + X_{COMP}D_{COMP} \qquad \textbf{Equation 4}$$

Substituting the values shown in Table 8 into Equation 4, the mole fraction, and thus the percentage of bound TSA in the complex was calculated and found to be a very significant 82%. Unfortunately, the exact molecular weight of the polymer is unknown,

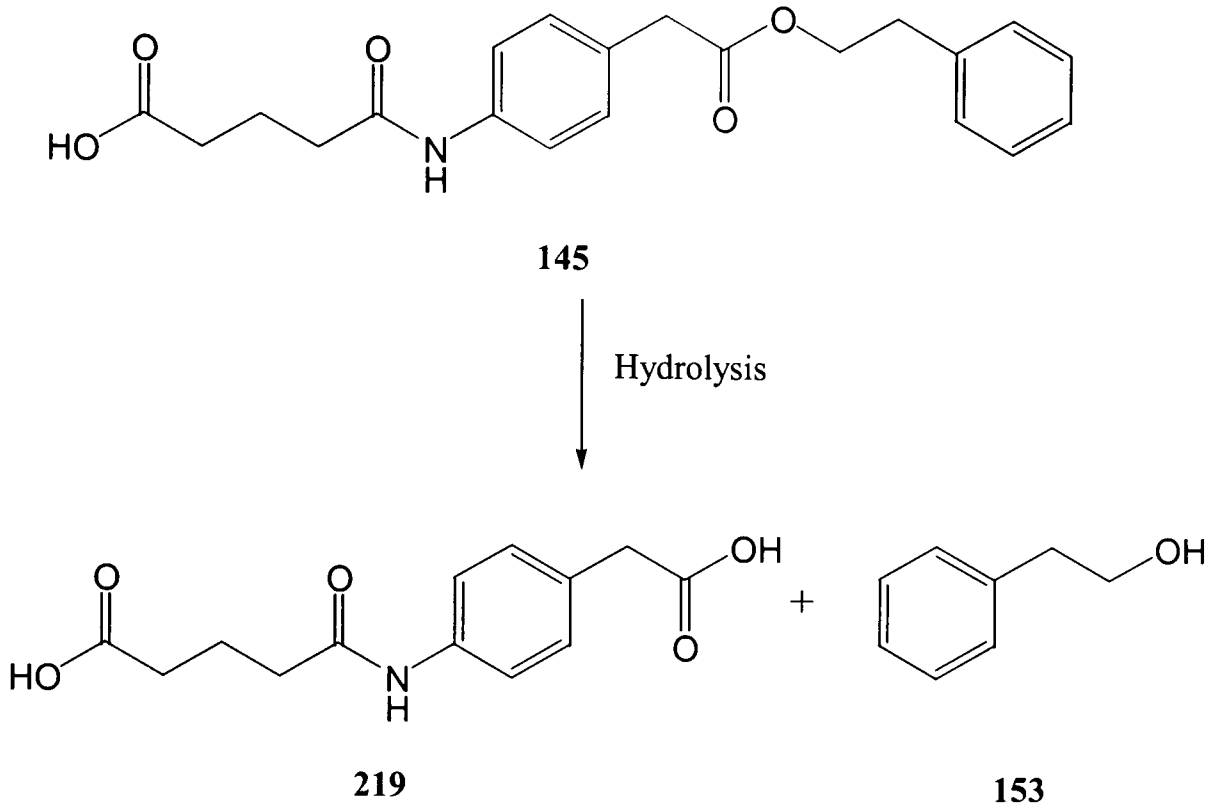
and hence it is not possible to know the concentration of the polymer in solution, and thus measure an exact binding constant K_a using Equation 6. However, an approximate value can be obtained, since the molecular weight of the starting polymer is known to be ~70,000. The concentration of the polymer solution is thus approximately 0.1mM, and this value can be substituted into Equation 6 to give an approximate value of K_a. The concentration of free and bound TSA can be calculated using the mole fraction found above.

$$K_a = \frac{[COMPLEX]}{[TSA][Polymer]} \qquad \textbf{Equation 6}$$

Substituting the values into Equation 6 gives a value for K_a of $45555LM^{-1}\pm20\%$. Although this figure is only an approximation, it indicates that the TSA has a strong affinity for the polymer. The results of these binding studies were encouraging, since if the polymer does indeed have a strong affinity for the TSA, then it should also bind the transition state in the ester hydrolysis, and hence have some catalytic affect on the reaction.

8.1 Hydrolysis Reactions of the Ester Substrate.

Before testing the polymers as catalysts for the hydrolysis of ester **145**, the rate of the control reaction, i.e. in the absence of any polymer had to be measured. The reactions were monitored by analytical HPLC. The hydrolysis of ester **145** yields two products, acid **219** and alcohol **153** (Scheme 65). Using a reverse phase column, and a solvent gradient elution system of acetonitrile and water, ester **145** and acid product **219** (Scheme 67) could be clearly separated with retention times of 18.01 and 11.9 minutes respectively, although phenethyl alcohol **153** could not be observed under these conditions.



Scheme 67

A 0.5mM solution of ester **145** in phosphate buffer pH7 was prepared, and stirred at room temperature. The reaction was monitored by HPLC after 10 minutes, and thereafter at thirty-minute intervals. After 24 hours, none of the peak corresponding to acid **219** was observed in the HPLC trace, and ^{1}H NMR of the sample confirmed that no reaction had occurred. The reaction was repeated, but on this occasion heated at 50°C for 24 hours. Again the HPLC trace showed only starting material.

Four polymers were now tested for their abilities to catalyse the hydrolysis of ester **145**. Since the control showed no reaction after 24 hours, any production of product during the same timescale could be seen as a significant rate increase. The four polymers used were the control polymer **204**, containing lysine residues only, polymer **218**, containing the Arg-Arg unit and lysine, polymer **217**, with lysine and histidine residues, and the vital polymer **215**, which contained all three units, Arg-Arg, lysine and histidine (Figure 43).

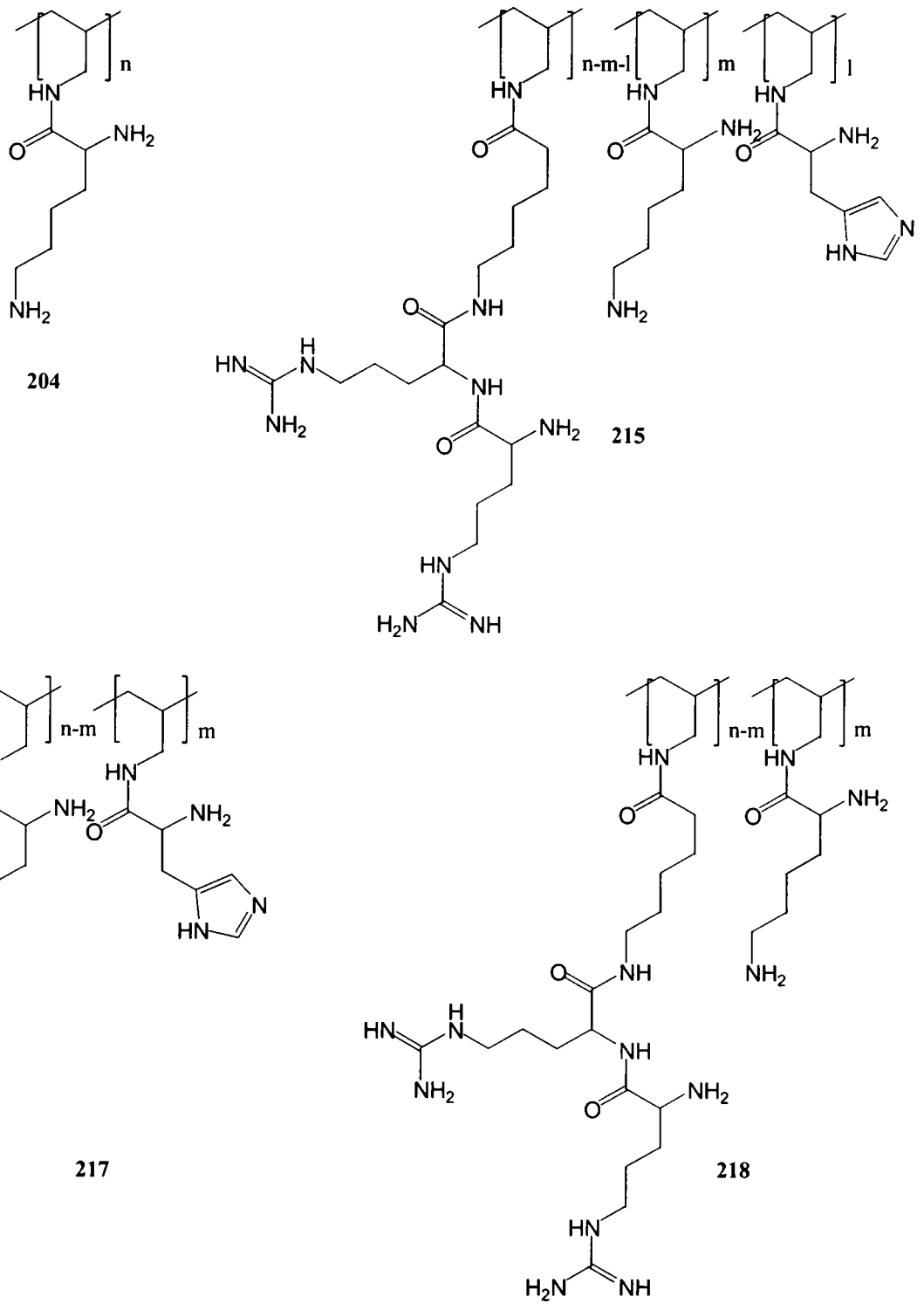


Figure 43

The conditions used for each polymer catalysed reaction were the same. A 0.5mM solution of **145** in phosphate buffer (pH7) was stirred with 15mg of polymer for 24 hours at room temperature. The reaction was monitored as described above using HPLC. Both the formation of acid **219**, and decrease in concentration of the starting material were monitored. In the presence of the control polymer **204**, after 24 hours, no significant reaction had occurred. The results for the other three polymers are summarised in Figures 44 and 45, showing the decrease in concentration of the ester **145** and the formation of the acid product **219** respectively.

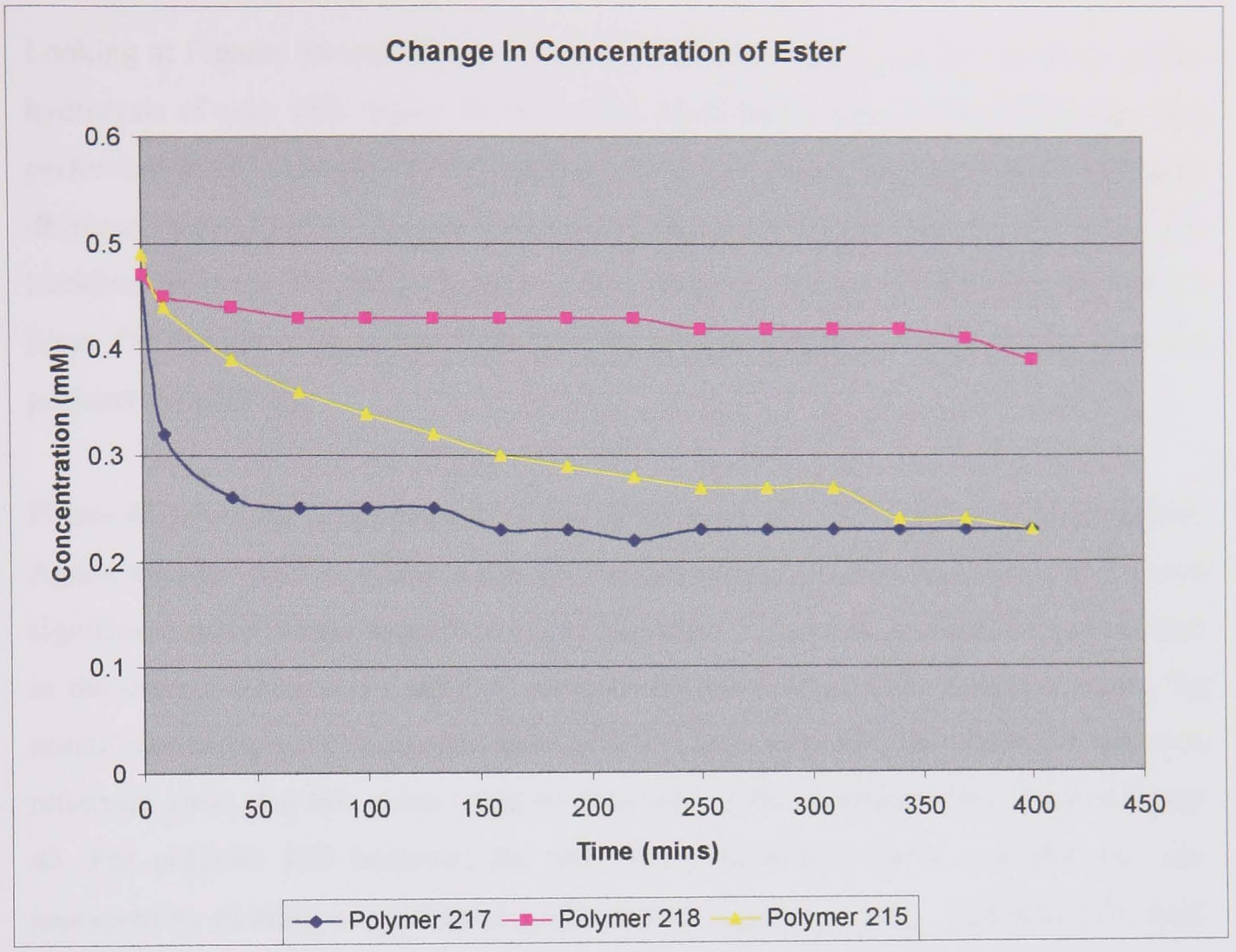


Figure 44

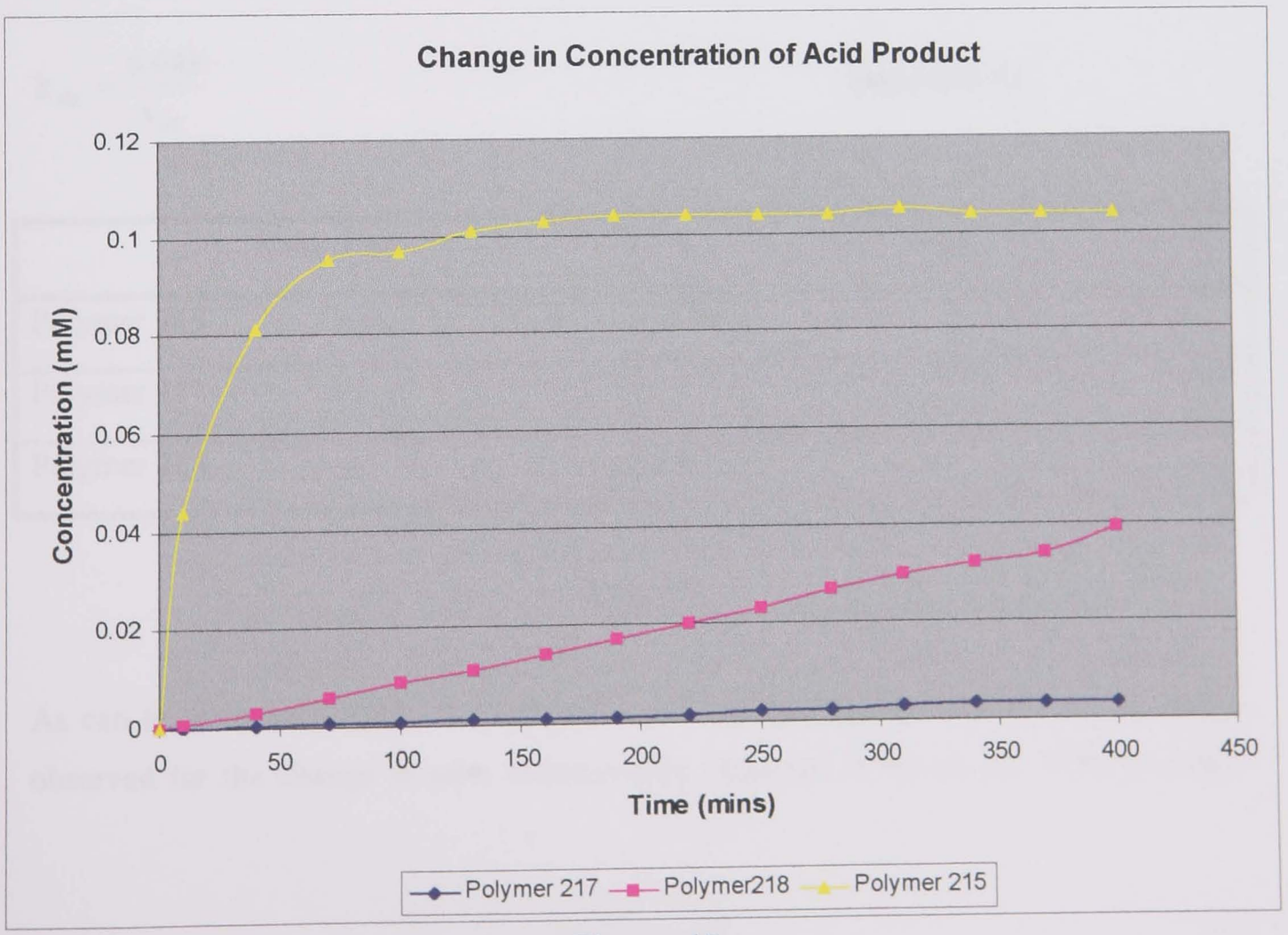


Figure 45

Looking at Figures 44 and 45 it is evident that all three polymers had an effect on the hydrolysis of ester **145**. Figure 44 shows that in all three cases, when the reaction was performed in the presence of the polymer, the concentration of ester **145** significantly decreased non-linearly with time. In the case of polymer **217**, containing only lysine and histidine residues, the concentration of the starting material had halved after only 1 hour. The results indicate that with all three polymers, binding of the substrate to the polymer is occurring.

Figure 45 shows the concentration of acid **219** formed during the course of the reaction. Again, all three polymers showed formation of product, which in itself can be seen as a significant achievement, when bearing in mind that no product whatsoever was formed in the control reaction. In order to compare the efficiency of the three polymers, the initial rate of reaction, k_{obs} was measured. For polymers **217** and **218**, the rate was relatively slow, and this value could be measured as the gradient of the slope in Figure 45. For polymer **215** however, the rate was considerably faster, and the rate was measured by plotting a logarithmic graph of time against $([Acid]_\infty - [Acid]_t)$. The half-life $t_{1/2}$ was read from the graph, and using Equation 12 the initial rate calculated. The rates for all three polymers are summarised in Table 9.

$$k_{obs} = \frac{0.643}{t_{1/2}} \qquad \textbf{Equation 12}$$

Polymer	k_{obs}
Polymer **215**	0.008
Polymer **217**	9×10^{-6}
Polymer **218**	9×10^{-5}

Table 9

As can be seen from Table 9, the relative rate constants give a reversal of the trends observed for the change in ester concentration observed in Figure 44. With an initial

rate constant of 9 x 10^{-5}, the reaction containing polymer **217**, containing lysine and histidine only showed the slowest rate of formation of product. Polymer **218**, containing the Arg-Arg receptor unit, showed a 10-fold rate acceleration over polymer **217**. Polymer **215**, containing both the receptor Arg-Arg unit, and histidine residues had the greatest effect on the rate of formation of product and showed a 900-fold rate acceleration over polymer **217**.

In order to obtain more information on how the reaction was proceeding, inhibition studies were carried out by performing the hydrolysis reactions in the presence of the TSA **146**. Should the catalysis be occurring due to binding and stabilisation of the transition state of the reaction, then the TSA should act as an inhibitor, and the rate of reaction should be significantly reduced. Reactions were carried out by stirring a solution with 0.5mM concentrations of both ester **145**, and TSA **146** in the presence of 15mg of polymer, and the reaction monitored as before. Unfortunately due to time constraints and a lack of available polymer, the inhibition studies were only carried out on polymers **217** and **218**. Figures 46 and 47 show the rate of formation of acid **219** for both polymers, in the presence and absence of TSA.

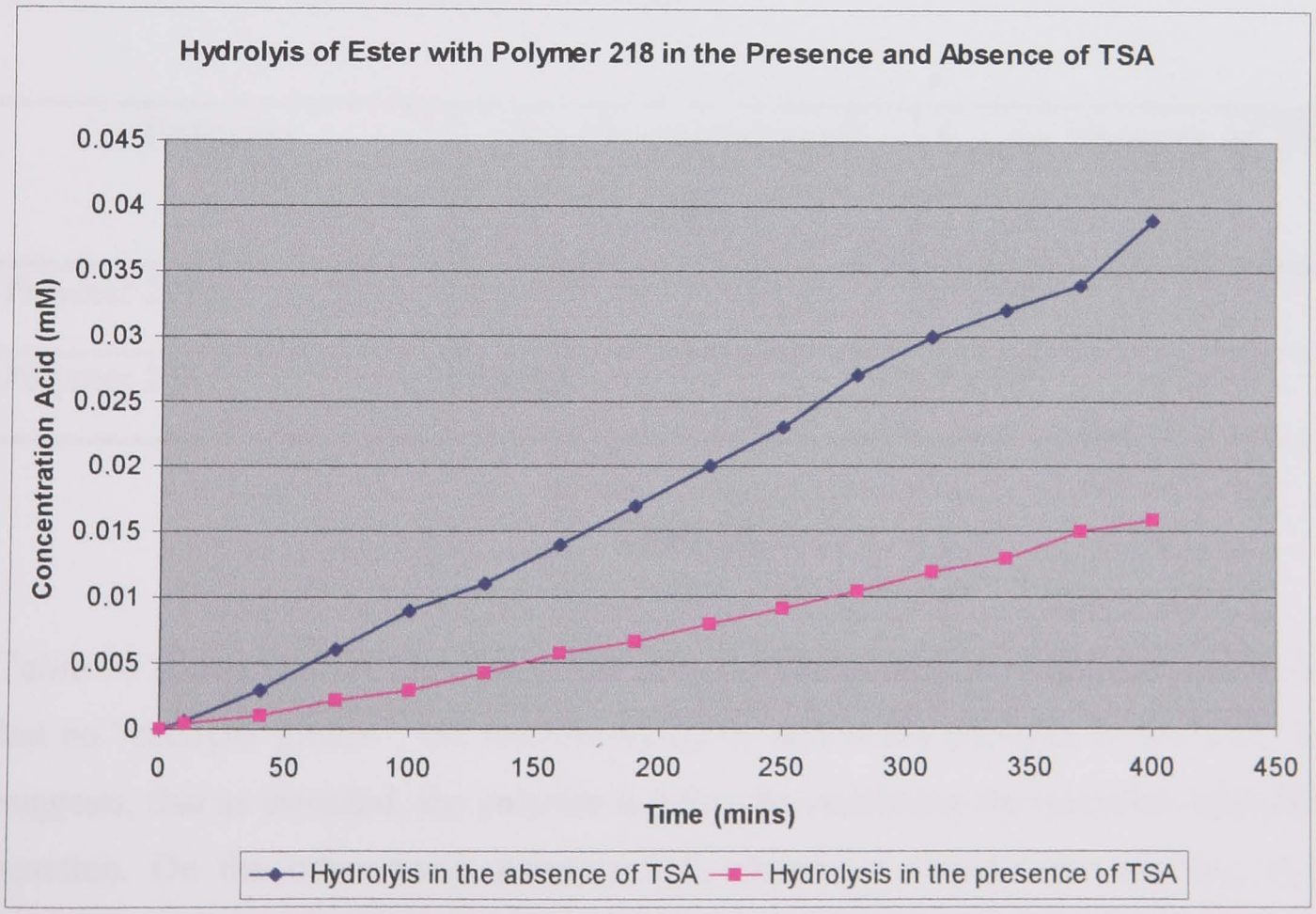


Figure 46

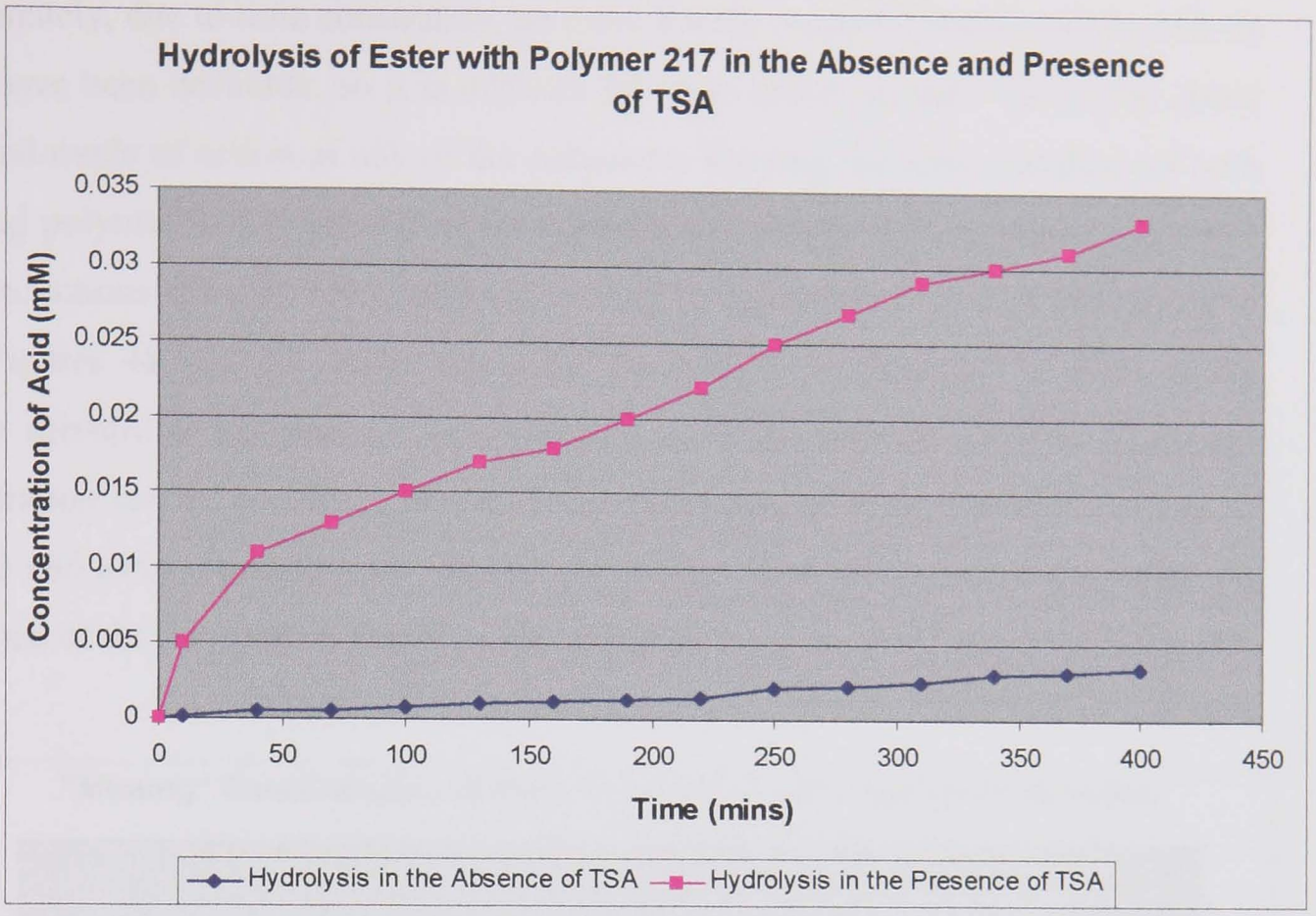


Figure 47

The initial rates of formation of product were measured in the same way as described above, and summarised in Table 10. For comparison, the rates calculated in the absence of TSA have also been included in the table.

Polymer	**k_{obs} in presence of TSA**	**k_{obs} in absence of TSA**
Polymer **217**	9×10^{-5}	9×10^{-6}
Polymer **218**	4×10^{-5}	9×10^{-5}

Table 10

Table 10 shows that for the case of the polymer containing the Arg-Arg receptor unit, but no "catalytic groups", the rate decreased by half in the presence of the TSA. This suggests, that as intended, the polymer is acting by stabilising the transition state of the reaction. On the other hand, polymer **217**, showed a 10-fold rate increase in the presence of the TSA. This result is interesting, and suggests a much more complex mechanism is involved in the catalysis.

Unfortunately, due to time constraints, no more kinetic studies could be performed, as would have been desirable, so it is difficult for us to make accurate statements about the actual mode of action of any of the polymers. Varying the concentrations of both ester and polymer would provide more information. However it is possible to make some deductions from the data collected. To aid in this discussion, it is interesting to study Figures 48 and 49, which show the "missing" concentration of ester in the reaction mixture in the absence and presence of TSA respectively. The "missing" concentration can be described as the initial concentration of ester minus the sum of the acid and ester concentrations at any given time. This data gives information on how much of the substrate is bound to the polymer.

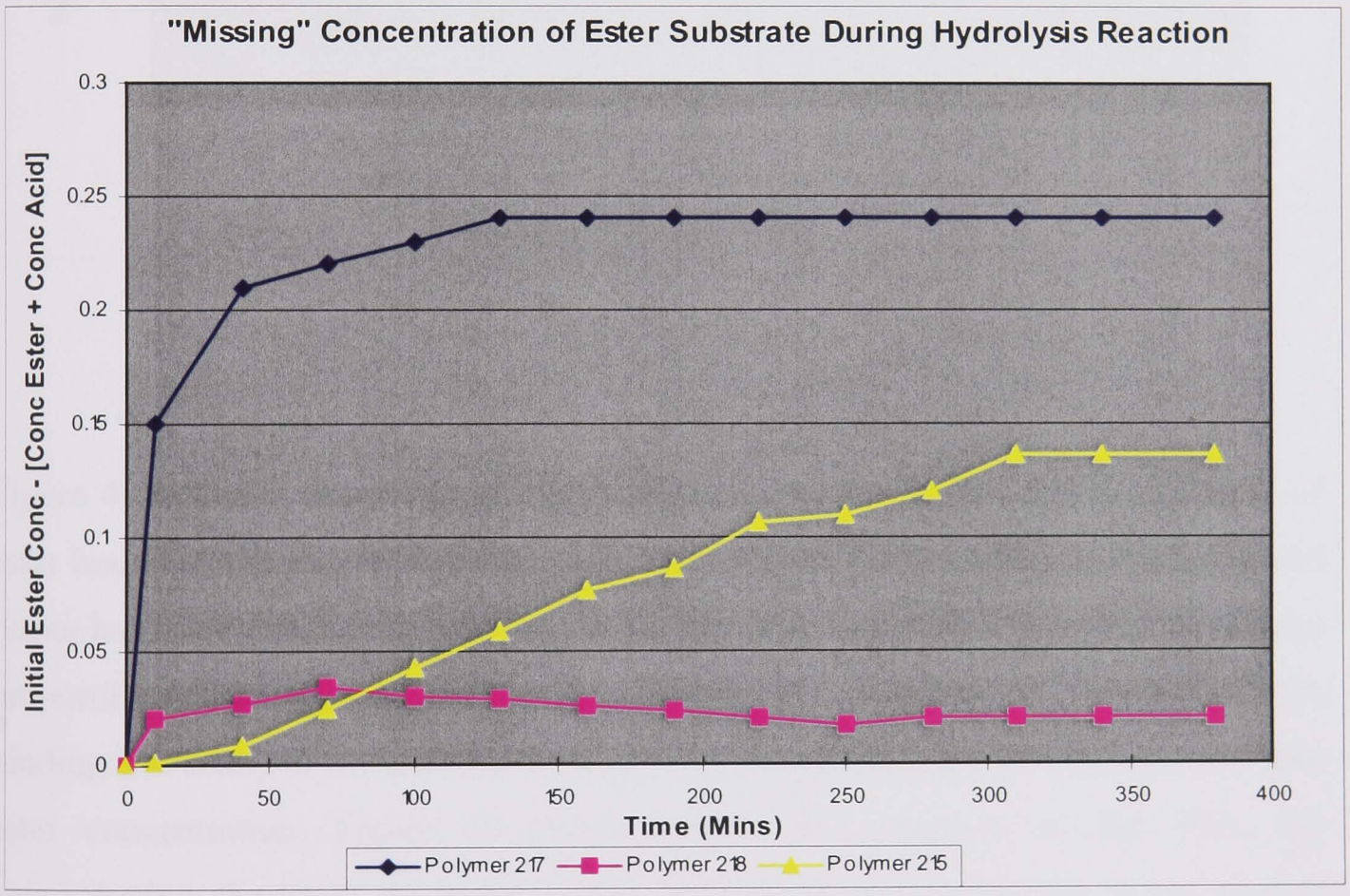


Figure 48

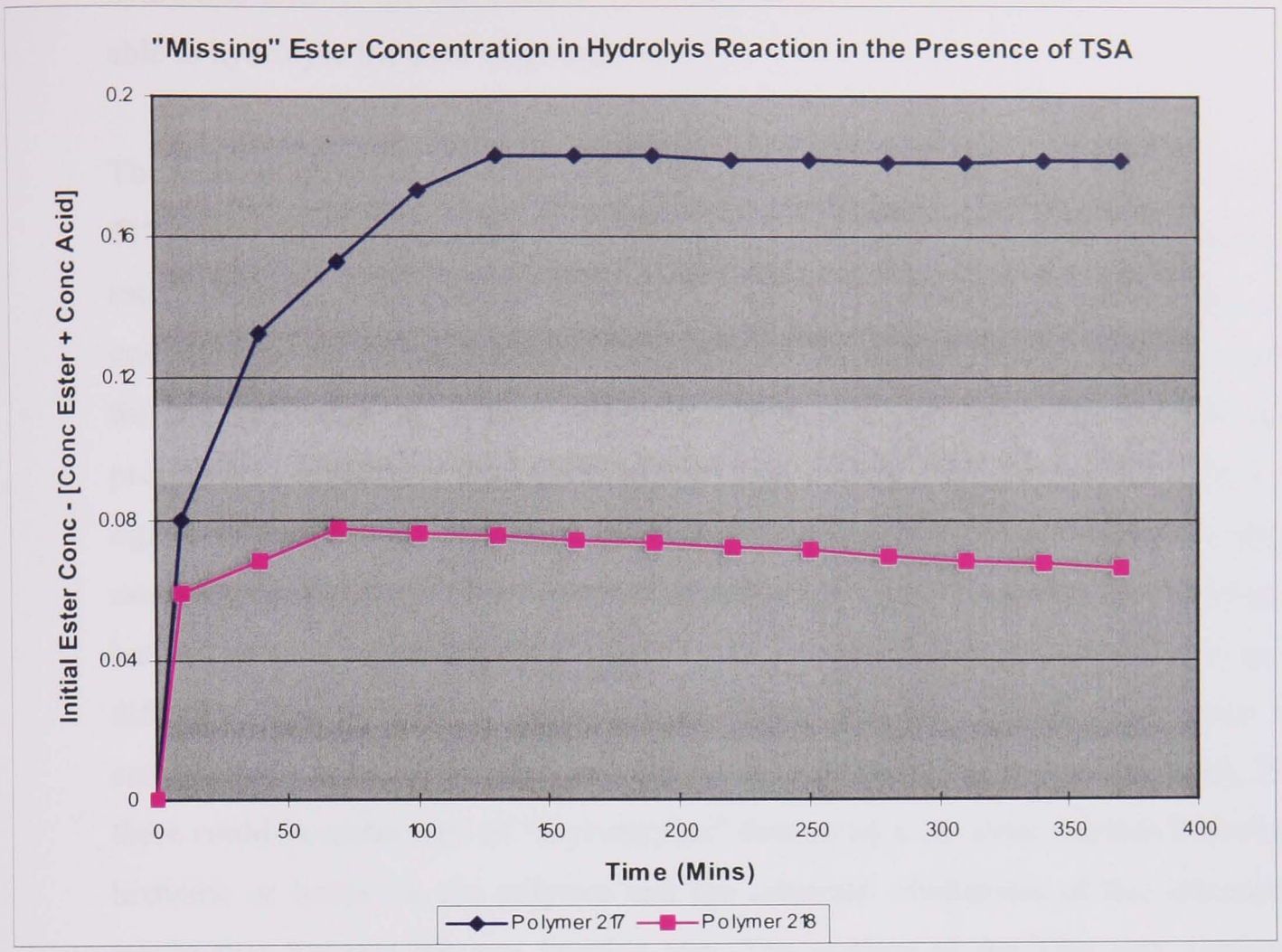


Figure 49

Figure 48 indicates that Polymer **218** has a low, but fairly constant concentration of ester bound to the polymer at any given time. Polymer **217** containing histidine and lysine has a very high "missing" concentration. This is a good indication that a large proportion of ester is strongly bound to the polymer. Polymer **215**, containing both binding and catalytic moieties shows a steadily increasing value of missing, or bound ester concentration. Figure 49 shows that in the presence of the TSA, the concentration shown in the graph varied little for Polymer **218**, but in the case of polymer **217** is almost halved.

Polymer **218**, containing the Arg-Arg receptor unit but no catalytic groups, shows a distinct rate increase over the control reaction, and very little of the ester substrate is bound to the polymer at any given time. The inhibition studies caused a 2-fold decrease in reaction rate, thus indicating that some aspect of the catalysis was caused by binding of a transition state. There is likely to be another factor in the catalysis by this polymer however. The guanidine groups of the arginine residues are protonated at

hydroxide ions. If the substrate is bound to the polymer then these hydroxide ions are able to hydrolyse the ester to give products.

The mode of action of the other two polymers is likely to be far more complex. Polymer **217** containing histidine and lysine residues apparently shows strong binding of the ester substrate, and a slow release of product. Approximately half of the initial concentration of ester is bound to the polymer at most times. When the TSA is added, the rate of release of product more than doubles. Several different theories can be proposed to explain the action of this polymer. Firstly, the substrate could bind very tightly to the polymer. However, this theory wouldn't explain the increase in reaction rate with the addition of TSA. Second, hydrolysis could be occurring at a binding site, but the product is binding very tightly to the polymer, either at a similar site, or at a different, second type of binding site. When the TSA is added, it may bind competitively with the product, and the rate of release of the acid is increased. Third, there could be some type of "acyl-enzyme" formed by a covalent reaction between the histidine or lysine on the polymer and the substrate. Hydrolysis of this intermediate would then become the rate limiting step. The binding of the TSA may change the binding, or conformation of binding sites and aid the hydrolysis of this complex. Without further studies however, it is impossible to accurately predict whether one of these, or another route is the actual mode of action.

What is clear is that the combination of both receptor site, and histidine in polymer **215** is required to form an efficient catalyst for the hydrolysis of ester **145**. One strange feature of this polymer is that as the reaction proceeds, the concentration of the product becomes constant, whilst the "missing" ester concentration steadily increases. This could be due to the system reaching equilibrium, so that the concentration of the product will no longer increase, but ester is still able to bind to the polymer, until the point where the polymer binding sites are saturated. The actual role of the histidine molecule is however unclear. In the active site of enzymes, histidine has a number of roles but most commonly acts as a general acid/base. It is also possible however that the histidine may, under certain conditions act as a nucleophile to form an acyl-imidazole intermediate, which should break down to give products. With the data in hand it is impossible to make a prediction about how the histidine is acting in this case.

In spite of the fact that the detailed mechanisms of the catalysis are not yet known, these preliminary results show significant rate enhancements, and thereby provide proof of the initial concept. Clearly, the modular assembly of designed receptor sites in combination with selection of catalytic groups represents a novel strategy for the assembly of artificial enzymes in a manner which also allows us to explore the various contributing factors which are responsible for enzymic catalysis.

9. Conclusions and Future Outlook.

9.0 Conclusions.

This thesis has described the development of a novel, "semi-rational" strategy for the preparation of artificial enzymes. Many of the previous attempts at developing artificial enzymes had achieved success in terms of catalysis, but in many instances, knowledge of the active site formed was limited. The initial concept for our approach was inspired by this fact. We wished to develop an approach in which it was possible to know which factors were responsible for catalysis, and also to be able to change them in order to make the catalyst more efficient. The idea was based on designing a known receptor site, and incorporating this onto a polymer backbone. The group responsible for binding and stabilising the transition state of the reaction was thus known. A range of active catalytic groups could then be grafted onto the polymer, and then the best system selected based on its ability to catalyse a given reaction, in our case ester hydrolysis.

The receptor molecule was chosen based on its ability to bind to a phosphonate TSA. The phosphonate contained three recognition points in the form of not only the phosphonic acid moiety, but also a carboxylic acid and amide linkage. A library of dipeptides were chosen as possible candidates since dipeptides possess appropriate functionality which could interact as required with the TSA. Binding interactions were studied using a combination of NMR and molecular modelling techniques. A novel method of studying small molecule-small molecule binding by measuring changes in diffusion coefficient by PFG-NMR, followed by molecular modelling to study the nature of the binding interactions identified the dipeptide Arg-Arg as a suitable receptor molecule to be incorporated into a polymer.

A range of amino acid residues, serine, histidine and cysteine were chosen to be incorporated alongside the receptor into a polymer as potential catalytically active groups. The choice of amino acids was based on knowledge of residues that play a vital role in the active site of proteases. Polyallylamine was selected as the polymer backbone since not only did it contain functionality suitable for coupling to the receptor and catalytic groups, but also because it was highly soluble and flexible. A flexible polymer

should be able to move freely in solution, and has the potential to exhibit "induced fit" as seen in natural enzymes, i.e. a change in conformation to bind a substrate.

Due to time limitations, only polymers containing the receptor and histidine could be prepared and tested for catalysis. A control reaction showed that the chosen ester did not hydrolyse under neutral conditions. A set of three polymers, one containing hisitidine, one containing the receptor, and one containing both were prepared and tested for catalysis of the hydrolysis of the chosen ester. In all three cases rate accelerations were observed. However, the polymer which contained both receptor and catalytic groups showed a 900-fold rate acceleration over the polymer containing just catalytic histidine residues, and a 90-fold rate acceleration over the polymer containing the receptor only.

This result neatly demonstrated the success of our initial concept. A receptor had been designed and incorporated into a polymer, and shown to catalyse the chosen ester hydrolysis. In addition, a suitable catalytic group had also been found. However, the most efficient catalysis was observed using a polymer which contained both receptor and catalytic residues. By clearly separating the catalytic groups and receptor, it was possible to investigate individually the effect and importance of each on the rate of reaction.

9.1 Future Outlook.

Our initial work in this area has successfully proved our proposed concept. However there is still plenty of scope for further studies.

With regard to the polymers already synthesised further kinetic studies should be performed which could give a greater understanding of the mechanism of catalysis. Firstly, studies varying the concentration of the substrate would enable us to calculate the maximum rate acceleration, and observe if the system follows Michaelis-Menton kinetics, operating by initial formation of a polymer-substrate complex.

Second, inhibition studies should be expanded upon. TSA inhibition studies should be performed on Polymer **215**, in a similar manner to polymers **217** and **218**. Also important is to study the effect of varying the concentration of the TSA on the rates of reaction. This information could provide an insight into the binding and mechanism of the polymers, especially polymer **217**.

Competition studies between the substrate that the system was designed for, and a simple ester, such as phenyl acetate, which contains no real "binding moieties" would also help in identifying the mode of binding. Clearly if the rate of hydrolysis of the substrate is increased relative to that of the competitive ester, then the binding of the substrate to the polymer, as predicted in Figure 28 is important. However, choice of ester here would be crucial, since the rate of hydrolysis of the substrate is exceedingly slow. The ester chosen would have to behave in a similar fashion.

Synthesis of the same polymers using the lower molecular weight polyallylamine should be performed. All of the catalytic features of the higher molecular weight species could still be incorporated into the polymer, but there would be added benefits, such as the increased solubility in a variety of solvents. This would enable the polymer to behave as a homogeneous catalyst in solution.

Synthesis of other polymers containing different "catalytic groups", and testing these as catalysts in the ester hydrolysis is also a priority. Due to the known role of serine and cysteine residues in natural enzymes, the use of these residues, both on their own, and in combination with histidine would be desirable.

Finally, investigations into finding a better "binding unit" could be performed. The binding studies described within this thesis were limited to nine dipeptides. However, there are clearly a vast number of possible peptides available, and the chances that we found the optimum binder are minimal. A more efficient binding assay, that utilized techniques such as UV and fluorescence could be set up, for initial studies, and once possible binders identified, the NMR diffusion technique could be used to identify the best binder.

The complexities involved in biomimetic catalysis are difficult to mimic using either conventional design, or more modern selection approaches. We have successfully demonstrated that the application of a "step by step" approach which combines both design and selection elements can be used, not only to synthesise an efficient catalyst, but also to provide at every step along the way detailed information about relative importance of the groups contributing to the molecular environment of the active site.

Chapter 3. Experimental.

1. General Experimental.

^{1}H NMR spectra were recorded at 300MHz on a Bruker AMX300 spectrometer, at 400MHz on a Bruker AMX400 spectrometer, or at 500MHz on a Bruker Avance 500 spectrometer. Chemical shifts (δ_H) are referenced to the residual solvent peak, except for spectra recorded in D_2O, in which case T.S.P was used as an internal standard. Chemical shifts are quoted in parts per million (ppm) using the following abbreviations: s, singlet; d, doublet; t, triplet; q, quartet; dd, double doublet; dt, doublet of triplets; m, multiplet; br, broad. The coupling constants (J) refer to vicinal $^{3}J_{HH}$ couplings unless otherwise stated, and are recorded in Hertz and quoted to the nearest 0.5Hz.

^{13}C NMR spectra were recorded at 75.4MHz on a Bruker AMX300 spectrometer, at 100.6MHz on a Bruker AMX400 spectrometer, or at 125.7MHz on a Bruker Avance 500 spectrometer. Chemical shifts (δ_C) are referenced to the residual solvent peak and are quoted to 0.1ppm. Carbon spectra assignments are supported by DEPT editing. Solid state ^{13}C NMR spectra were recorded at 75.4MHz on a Bruker MSL300 spectrometer.

^{31}P NMR were recorded at 121.4MHz on a Bruker AMX300 spectrometer. Chemical shifts (δ_P) are referenced to 85% aq. H_3PO_4, and are quoted to 0.1ppm.

2D- NOESY NMR spectra were carried out on a Bruker Avance 500 spectrometer. PFG Diffusion Ordered SpectroscopY (DOSY) NMR experiments were carried out on a Bruker Avance 500 NMR spectrometer equipped with a Bruker GAB-type gradient accessory.

Melting points were determined using a Reichart Hotstage or Electrothermal 9100 instrument and are uncorrected.

Infrared spectra were recorded on a Perkin-Elmer 1600 Fourier transform spectrometer, and were recorded either as thin films (NaCl), as KBr disks, or as $CHCl_3$ solutions (NaCl). Absorption maxima are reported in wavenumbers (cm^{-1}) using the following abbreviations: w, weak; m, medium; s, strong; br, broad. Only selected absorbances are

reported. Polymer solid state IR were performed using a microscope IR at Imperial College, London.

Mass spectra were recorded under either fast atom bombardment (FAB), or atmospheric pressure chemical ionisation (APCI) conditions by the ULIRS mass spectrometry service at the School of Pharmacy, University College London. Only molecular ions, or fragments from molecular ions are reported. High resolution mass spectra are accurate to ±10ppm.

Microanalyses were performed in the University College London Chemistry Department.

DMF was dried with $MgSO_4$, and then distilled from silica gel and stored over Linde Type 4Å molecular sieves. Diisopropylethylamine (DIPEA) was distilled from, and stored over KOH. THF was freshly distilled from sodium-benzophenone under N_2. Benzene was distilled from sodium, and DCM was distilled from calcium hydride. Methanol was distilled using magnesium turnings and iodine. Triethylamine was distilled from P_2O_5 and stored over KOH. *N*-Methylmorpholine was distilled from calcium hydride and stored over Linde Type 4Å molecular sieves. *Tert*-Butylamine was distilled neat immediately prior to use.

With the exception of FmocLysine(Fmoc)OH which was synthesised in house, all protected amino acids, and dipeptides were purchased from BACHEM and used as bought.

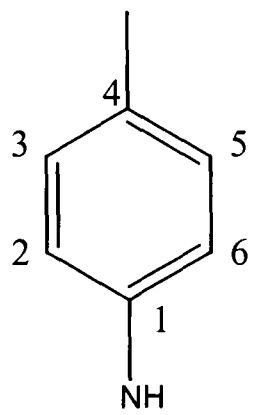
Analytical HPLC was carried out on a Shimadzu dual pump LC-10AS with a Shimadzu SPD-6A UV/VIS variable wavelength detector, and Shimadzu C-R6A chromatapac integrator, and a reverse phase C18 column (4.6 by 250mm) from HiChrom. Solvents were HPLC grade and degassed prior to use.

Analytical thin layer chromatography was performed on aluminium sheets pre-coated with Merck silica gel 60 F_{254}, and visualised with ultraviolet light (254nm), plus either basic potassium permanganate, acidic vanillin, or acidic anisaldehyde solution. Flash chromatography was performed using BDH silica gel (40-60μm).

All glassware was oven dried and cooled under a flow of nitrogen. All reactions requiring dry solvents were carried out under an atmosphere of nitrogen.

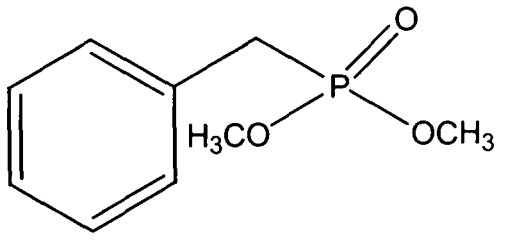
2. Synthesis of Phosphonate Transition State Analogue and Ester Substrate

For consistency when assigning NMR spectra, the following numbering system is used for assignment purposes only in the *para*-substituted benzene rings:



2.0 Synthesis of the Transition State Analogue (TSA).

Synthesis of Benzyl Phosphonic Acid Dimethyl Ester (149)[158].

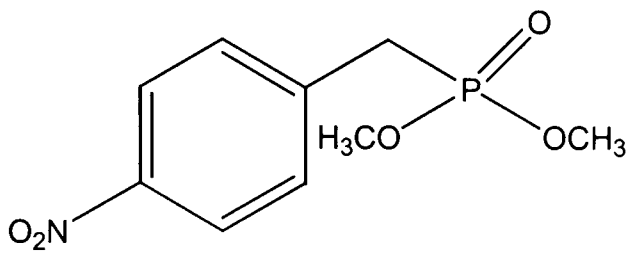


149

A stirred solution of trimethyl phosphite (195ml, 1.65mol) and benzyl bromide (27.82ml, 0.234mol) in benzene (100ml) was heated at reflux for 19 hours. The solvent was removed *in vacuo*, and the excess trimethyl phosphite removed by distillation to yield a colourless oil. This was further purified by fractional distillation (135-138°C, 3.5mmHg) to afford benzyl phosphonic acid dimethyl ester **149** as a clear colourless oil (27g, 93%); $\mathbf{R_f}$ 0.36 (SiO_2; EtOAc); **^{1}H NMR** (500MHz, $CDCl_3$) δ_H/ppm 3.13 (d, $^2J_{PH}$=21.5, ArC$\underline{H}_2$P), 3.63 (d, $^3J_{POCH}$=11.0, 6H, 2 x OC$\underline{H}_3$), 7.21-7.36 (m, 5H, 5 x Ar$\underline{H}$); **^{13}C NMR** (100.6MHz, $CDCl_3$) δ_C/ppm 32.8 (d, J_{CP}=138.0, Ar$\underline{C}H_2$P), 52.7 (2 x O$\underline{C}H_3$), 126.9 ($\underline{C}$(1)), 128.6 ($\underline{C}$(3), $\underline{C}$(5)), 129.6 ($\underline{C}$(2), $\underline{C}$(6)), 131.5 (1 x quaternary, $\underline{C}$(4)); ν_{max} (thin film/cm^{-1}) 584(m), 700(m), 756(m), 832(s), 916(m), 1026(s, P-OMe), 1186(m), 1250(s, P=O), 1406(m), 1466(m), 1496(m), 1603(m), 2844(m), 2891(m), 2957(s),

2970(m), 3170(m), 3400(s), 3631(s); **m/z** (FAB) 201 (100%, $[M+H]^+$); **HRMS** found 201.0675, $[M+H]^+$ ($C_9H_{14}O_3P$) requires 201.0681.

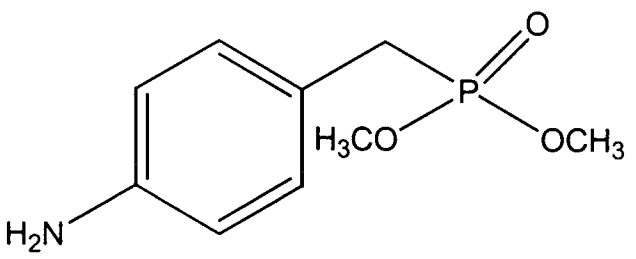
Synthesis of (4-Nitro-Benzyl)-Phosphonic Acid Dimethyl Ester (156)[159].



156

Nitric acid (70% w/v, 24.1ml) and sulphuric acid (98% w/v, 24.1ml) were mixed together with external cooling. When the temperature reached 0°C, benzyl phosphonic acid dimethyl ester **149** (19.2g, 0.096mol) was added dropwise over 45mins. The mixture was stirred at 0°C for 1 hour, after which time the reaction was quenched by addition to ice H_2O (600ml). The product was extracted with toluene (4 x 150ml), the organic layers were combined, washed with aq. $NaHCO_3$ (sat., 2 x 100ml), aq. NaCl (sat., 1 x 100ml), and dried over Na_2SO_4. The solvent was removed *in vacuo* to yield (4-nitro-benzyl)-phosphonic acid dimethyl ester **156** as a pale yellow solid (10.13g, 79%); **m.p.** 68-70°C (lit. 75°C[159]); $\mathbf{R_f}$ 0.22 (SiO_2; EtOAc); ^{1}H NMR showed a 9:1 ratio of *para*:*ortho* isomers; Data for *para* isomer only; **^{1}H NMR** (500MHz, $CDCl_3$) δ_H/ppm 3.19 (d, $^2J_{PH}$=22.5, 2H, ArC$\underline{H}_2$P), 3.62 (d, $^3J_{POCH}$=11.0, 6H, 2 x OC$\underline{H}_3$), 7.39 (dd, $^4J_{PH}$=2.5, J=8.5, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 8.08 (d, J=8.5, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 33.3 (d, J_{CP}=138.0, Ar$\underline{C}H_2$P), 53.4, 53.5 (2 x O$\underline{C}H_3$), 124.2 ($\underline{C}$(2), $\underline{C}$(6)), 131.0 ($\underline{C}$(3), $\underline{C}$(5)), 139.8, 147.4 (2 x quaternary, $\underline{C}$(1), $\underline{C}$(4)); ν_{max} ($CHCl_3$/cm^{-1}) 549(m), 689(m), 740(m), 829(m), 864(s), 1032(s), 1053(s, P-OMe), 1186(m), 1249(s, P=O), 1347(s, C-NO_2), 1403(w), 1452(w), 1487(w), 1515(s, C-NO_2), 1599(m), 2454(w), 2853(w), 2851(m), 3000(s), 3414(m, br); **m/z** (FAB) 246 (100%, $[M+H]^+$); **HRMS** found 246.0540, $[M+H]^+$ ($C_9H_{13}NO_5P$) requires 246.0531.

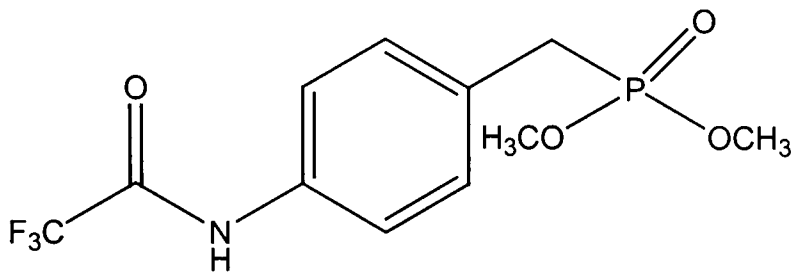
Synthesis of (4-Amino-Benzyl)-Phosphonic Acid Dimethyl Ester (150)[160].



150

To a solution of (4-nitro-benzyl)-phosphonic acid dimethyl ester **156** (15g, 0.061mol) in MeOH (200ml) was added 10% Pd/C (1.5g, cat.). The mixture was placed in a Parrs hydrogenation equipment under a hydrogen pressure of 50 p.s.i. and shaken for 24 hours. The catalyst was removed by filtration, and the solvent removed *in vacuo*. The residue was dissolved in aq. HCl (2M, 100ml), and any unreacted starting material extracted with toluene (2 x 90ml). The aqueous layer was adjusted to pH9 using $NaHCO_3$, and the product extracted into DCM (4 x 80ml). The organic layers were combined, washed with H_2O (1 x 80ml), aq. NaCl (sat., 1 x 80ml), dried over Na_2SO_4, and concentrated to ~20ml. On standing crystallisation occurred to afford pink crystals, which were collected by filtration to yield (4-amino-benzyl)-phosphonic acid dimethyl ester **150** (6.95g, 53%); **m.p.** 105-107°C (lit. 103°C[160]); $\mathbf{R_f}$ 0.34 (SiO_2; EtOAc:MeOH; 9:1); **^{1}H NMR** (500MHz, $CDCl_3$) δ_H/ppm 3.00 (d, $^2J_{PH}$=21.0, 2H, ArC$\underline{H}_2$P), 3.59 (d, $^3J_{POCH}$=11.0, 6H, 2 x OC$\underline{H}_3$), 6.57 (d, J=8.0, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$), 7.01 (dd, $^4J_{PH}$=2.5, J=8.0, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 32.7 (d, J_{CP}=136.0, Ar$\underline{C}H_2$P), 53.7, 53.8 (2 x O$\underline{C}H_3$), 116.1, ($\underline{C}$(2), $\underline{C}$(6)), 121.3 (1 x quaternary, $\underline{C}$(4)), 131.3, 131.4 ($\underline{C}$(3), $\underline{C}$(5)), 146.4 (1 x quaternary, $\underline{C}$(1)); ν_{max} ($CHCl_3$/cm^{-1}) 563(m), 661(m), 696(m), 815(m), 857(s), 1052(s, P-OMe), 1130(w), 1179(s), 1242(s), 1263(s, P=O), 1403(m), 1452(m), 1508(s, Ar), 1621(s, Ar), 2853(w), 2951(m), 3001(s), 3386(m, -NH_2), 3449(m, -NH_2); **m/z** (FAB) 216 (70%, $[M+H]^+$); **HRMS** found 216.0776, $[M+H]^+$ ($C_9H_{15}NO_3P$) requires 216.0790.

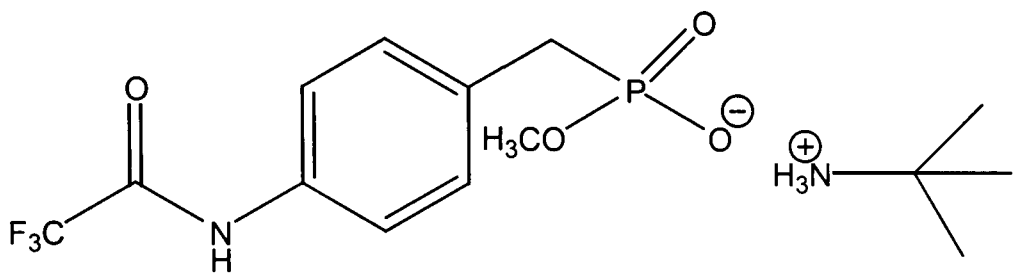
Synthesis of [4-(2,2,2-Trifluoroacetylamino)-Benzyl]-Phosphonic Acid Dimethyl Ester (151).



151

To a stirred solution of (4-amino-benzyl)-phosphonic acid dimethyl ester **150** (4.04g, 0.019mol) in anhydrous THF (40ml) under N_2 was added trifluoroacetic anhydride (2.71ml, 0.019mol) and *N*-methylmorpholine (2.11ml, 0.019mol). The solution was stirred at r.t. for 1.5 hours. The solvent was removed *in vacuo*, and the residue dissolved in EtOAc (100ml). The solution was washed with aq. $KHSO_4$ (1N, 4 x 20ml), aq. $NaHCO_3$ (5%, 2 x 20ml), aq. NaCl (sat., 1 x 20ml), and dried over Na_2SO_4. The solvent was concentrated *in vacuo* until precipitation occurred. The solid was collected by filtration to yield [4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid dimethyl ester **151** as a white crystalline solid (4.69g, 80%); **m.p.** 130-132°C; $\mathbf{R_f}$ 0.47 (SiO_2; EtOAc:MeOH; 9:1); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 3.16 (d, $^2J_{PH}$=21.5, 2H, ArC$\underline{H}_2$P), 3.75 (d, $^3J_{POCH}$=11.0, 2H, 2 x OC$\underline{H}_3$), 7.13 (dd, $^4J_{PH}$=2.5, J=8.5, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.51 (d, J=8.5, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$), 10.79 (s, 1H, F_3CC(O)N$\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 32.4 (d, J_{CP}=139.5, Ar$\underline{C}H_2$P), 53.4, 53.5 (2 x O$\underline{C}H_3$), 121.7 ($\underline{C}$(2), $\underline{C}$(6)), 128.2 ($\underline{C}F_3$C(O)NH), 130.6 ($\underline{C}$(3), $\underline{C}$(5)), 135.8, 135.9 (2 x quaternary, $\underline{C}$(1), $\underline{C}$(4)), 155.9 ($CF_3\underline{C}$(O)NH); **^{31}P NMR** (121.4MHz, $CDCl_3$) δ_P/ppm 30.1; ν_{max} ($CHCl_3$/cm^{-1}) 559(m), 725(m), 787(w), 822(m), 866(m), 902(m), 1099(s, C-F), 1062(s, P-OMe), 1157(s, C-F), 1195(s, C-F), 1259(s, P=O), 1287(m), 1421(m), 1516(m, C(O)NH), 1555(m), 1606(m), 1718(s, C=O), 2951(m), 3000(m), 3042(s), 3197(m), 3253(m, C(O)NH), 3421(m, C(O)NH); **m/z** (FAB) 312 (100%, $[M+H]^+$); **HRMS** found 312.0620, $[M+H]^+$ ($C_{11}H_{14}F_3NO_4P$) requires 312.0613.

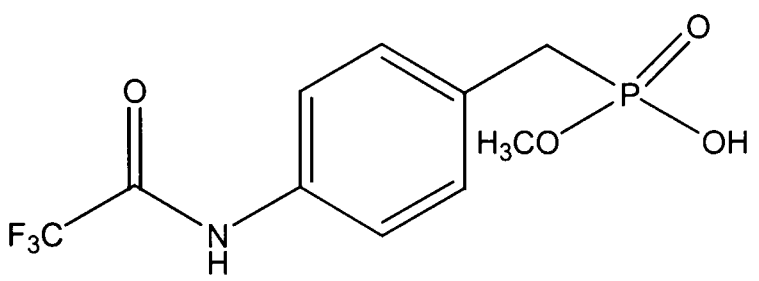
Synthesis of [4-(2,2,2-Trifluoroacetylamino)-Benzyl]-Phosphonic Acid *tert*-Butylamine Salt Methyl Ester (158).



158

[4-(2,2,2-Trifluoroacetylamino)-benzyl]-phosphonic acid dimethyl ester **151** (3.08g, 9.91mmol) was suspended in *tert*-butylamine (30ml) and the mixture refluxed for 7 days. The reaction was concentrated *in vacuo*, and the residue dissolved in H_2O, and lypholised to yield [4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid *tert*-butylamine salt methyl ester **158** as a white solid (3.4g, 92%); $\mathbf{R_f}$ 0.21 (SiO_2; $CHCl_3$:MeOH; 3:1); **^{1}H NMR** (400MHz, D_2O) δ_H/ppm 1.12 (s, 9H, $(C\underline{H}_3)_3CNH_3$), 3.33 (d, $^2J_{PH}$=21.5, 2H, $ArC\underline{H}_2P$), 4.67 (d, $^3J_{POCH}$=11.0, 3H, $OC\underline{H}_3$), 6.94 (d, J=7.5, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.08 (d, J=7.5, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (100.6MHz, D_2O) δ_C/ppm 27.9 ($(\underline{C}H_3)_3CNH_3$), 33.8 (d, J_{CP}=130.0, $Ar\underline{C}H_2P$), 52.3 ($O\underline{C}H_3$), 117.1, 117.5 (2 x quaternary), 123.8, 123.9 (C(2), C(6)), 130.9, 131.0 (C(3), C(5)), 131.7, 141.9 (2 x quaternary); **^{31}P NMR** (121.4MHz, D_2O) δ_P/ppm 22.4.

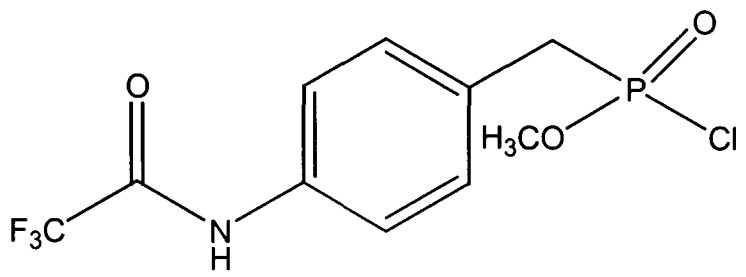
Synthesis of [4-(2,2,2-Trifluoroacetylamino)-Benzyl]-Phosphonic Acid Methyl Ester (159).



159

tert-Butylamine salt **158** (3g, 8mmol) was dissolved in MeOH (50ml) and stirred with cation exchange resin Amberlite IR-120 (plus) for 4 hours. The resin was removed by filtration and the solvent removed *in vacuo* to afford [4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid methyl ester **159** as a cream coloured solid (1.75g, 68%); **m.p.** 206-208°C; **^{1}H NMR** (500MHz, CD_3OD) δ_H/ppm 3.18 (d, $^2J_{PH}$=21.5, 2H, ArC$\underline{H}_2$P), 3.65 (d, $^3J_{POCH}$=11.0, 3H, OC$\underline{H}_3$), 7.33 (dd, $^4J_{PH}$=2.5, J=8.5, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.58 (d, J=8.5, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (75.4MHz, CD_3OD) δ_C/ppm 33.9 (d, $^2J_{PH}$=139.5, Ar$\underline{C}H_2$P), 53.3 (O$\underline{C}H_3$), 115.9, 119.8 (2 x quaternary), 122.6 ($\underline{C}$(2), $\underline{C}$(6)), 131.6 ($\underline{C}$(3), $\underline{C}$(5)), 136.9 (1 x quaternary), 157.4 ($CF_3\underline{C}$(O)NH); ν_{max} (KBr/cm^{-1}) 504(m), 556(m), 699(m), 835(s), 1046(m, P-OMe), 1162(s, C-F), 1244(m, P=O), 1285(m), 1346(w), 1414(m), 1462(m), 1516(m), 1543(m), 1605(m), 1721(s, C=O), 2858(w), 2913(w), 2960(m), 3083(m), 3137(m), 3206(m, C(O)NH), 3315(m, C(O)NH); **m/z** (FAB) 298 (80%, $[M]^+$); 136 (100%) **HRMS** found 298.0443, $[M]^+$ ($C_{10}H_{11}F_3NO_4P$) requires 298.0456.

Synthesis of [4-(2,2,2-Trifluoroacetylamino)-Benzyl]-Phosphonic Acid Chloride Methyl Ester (152).



152

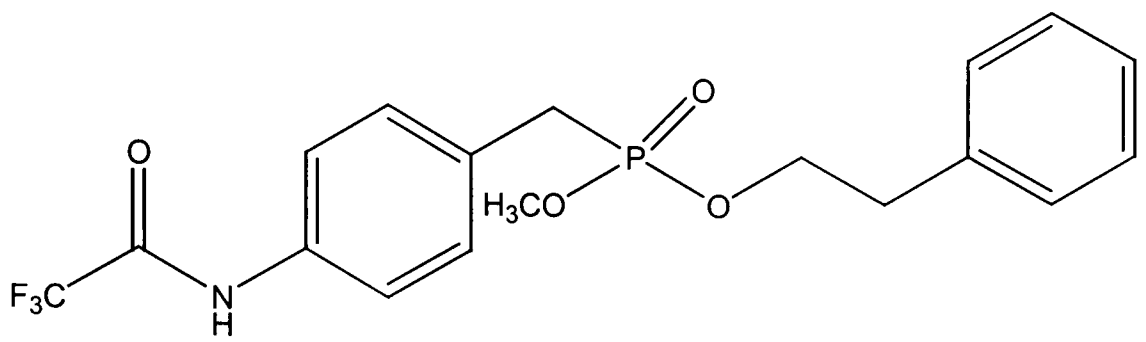
Method 1.

Phosphorus pentachloride (0.56g, 2.6mmol) was added to a stirred solution of [4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid dimethyl ester **151** (0.5g, 1.7mmol) in $CHCl_3$ (5ml) under N_2 and heated at 40°C for 16 hours. SO_2 gas was bubbled through the solution for 10 mins. to remove any excess phosphorus pentachloride. The solvent was removed *in vacuo* to afford [4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid chloride methyl ester **152** as a brown solid (0.502g, 98% crude). Analytical data as below.

Method 2.

[4-(2,2,2-Trifluoroacetylamino)-benzyl]-phosphonic acid methyl ester **159** (4g, 0.013mol) was suspended in anhydrous DCM (50ml) under N_2. Oxalyl chloride (1.76ml, 0.02mol) was added followed by the dropwise addition of DMF (51µl, 0.7mmol). The reaction was stirred for 24 hours at r.t. After this time the solution had become homogeneous. The solvent was removed *in vacuo*, and the solid washed with toluene (1 x 20ml), to yield [4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid chloride methyl ester **152** as a cream coloured solid (4.12g, 99%); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 3.58 (dd, $^2J_{PH}$=21.0, J=7.5, 2H, ArC$\underline{H}_2$P), 3.95 (d, $^3J_{POCH}$=13.5, 3H, OC$\underline{H}_3$), 7.28 (dd, $^4J_{PH}$=3.0, J=8.0, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.62 (d, J=8.0, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$), 9.73 (s, 1H, $CF_3C(O)N\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 40.4 (d, J_{CP}=139.0, Ar$\underline{C}H_2$P), 53.8 (O$\underline{C}H_3$), 114.4, 118.2 (2 x quaternary), 121.6 ($\underline{C}$(3), $\underline{C}$(5)), 131.1 ($\underline{C}$(2), $\underline{C}$(6)), 136.3 (1 x quaternary), 155.5 ($CF_3\underline{C}(O)NH$); **^{31}P NMR** (121.4MHz, $CDCl_3$) δ_P/ppm 42.1; **m/z** (FAB) 315, 317 (100%, $[M]^+$).

Synthesis of [4-(2,2,2-Trifluoroacetylamino)-Benzyl]-Phosphonic Acid Methyl Ester Phenethyl Ester (154).

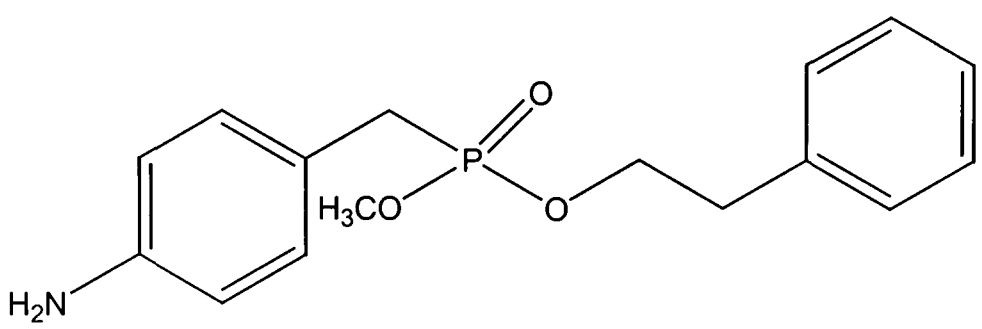


154

[4-(2,2,2-Trifluoroacetylamino)-benzyl]-phosphonic acid chloride methyl ester **152** (4.1g, 0.013mol) was dissolved in anhydrous THF (60ml) and added dropwise to a stirred solution of 2-phenylethanol (1.55ml, 0.013mol) and triethylamine (3.6ml, 0.026mol) in anhydrous THF (20ml) under N_2. The solution was stirred for 24 hours at r.t. After this time triethylamine hydrochloride had formed as a white precipitate. The precipitate was removed by filtration, and the solvent evaporated. The residue was dissolved in EtOAc (60ml) and washed with aq. $NaHCO_3$ (5%, 2 x 30ml), aq. $KHSO_4$ (5%, 2 x 40ml), aq. NaCl (sat., 1 x 50ml), and dried over Na_2SO_4. The solvent was removed *in vacuo* to afford a yellow oil that was purified by flash chromatography (SiO_2; EtOAc) to yield [4-(2,2,2 trifluoroacetylamino)-benzyl]-phosphonic acid methyl ester phenethyl ester **154** as a yellow solid (4.11g, 79%); **m.p.** 135-137°; $\mathbf{R_f}$ 0.34 (SiO_2; EtOAc); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 2.96 (t, *J*=7.0, 2H, $POCH_2C\underline{H}_2Ar$), 3.10 (d, $^2J_{PH}$=21.5, 2H, $ArC\underline{H}_2P$), 3.62 (d, $^3J_{POCH}$=11.0, 3H, $OC\underline{H}_3$), 4.14-4.27 (m, 2H, $POC\underline{H}_2CH_2Ar$), 7.14 (dd, $^4J_{PH}$=2.5, *J*=8.0, 2H, C(3)H, C(5)H), 7.21 (d, *J*=7.0, 2H, *ortho*-ArH), 7.30 (t, *J*=7.0, 1H, *para*-ArH), 7.32 (t, *J*=7.0, 2H, *meta*-ArH), 7.55 (d, *J*=8.0, 2H, C(2)H, C(6)H); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 32.7 (d, J_{CP}=138.5, $Ar\underline{C}H_2P$), 37.3 ($POCH_2\underline{C}H_2Ar$), 53.1 ($O\underline{C}H_3$), 67.3 ($PO\underline{C}H_2CH_2Ar$), 119.2 (1 x quaternary, $\underline{C}F_3C(O)NH$), 121.7 (C(2), C(6)), 127.2 (*para*-ArC) 128.9, 129.3, (2 x *ortho*-ArC, 2 x *meta*-ArC), 130.6 (C(3), C(5)), 128.1, 136.2, 137.5 (3 x quaternary, ArC), 155.9 ($CF_3\underline{C}(O)NH$); ν_{max} ($CHCl_3$/cm^{-1}) 557(m), 650(m), 721(s), 769(s), 1018(m), 1061(m, P-OR), 1157(m, C-F), 1247(s, P=O), 1287(m), 1380(m), 1462(s), 1515(m), 1551(m), 1610(m), 1723(s, C=O), 3017(m), 3461(s, C(O)NH), 3593(s,

C(O)NH); **m/z** (APCI$^+$) 402 (100%, $[M+H]^+$); **HRMS** found 402.1082, $[M+H]^+$ ($C_{18}H_{20}F_3NO_4P$) requires 402.1090.

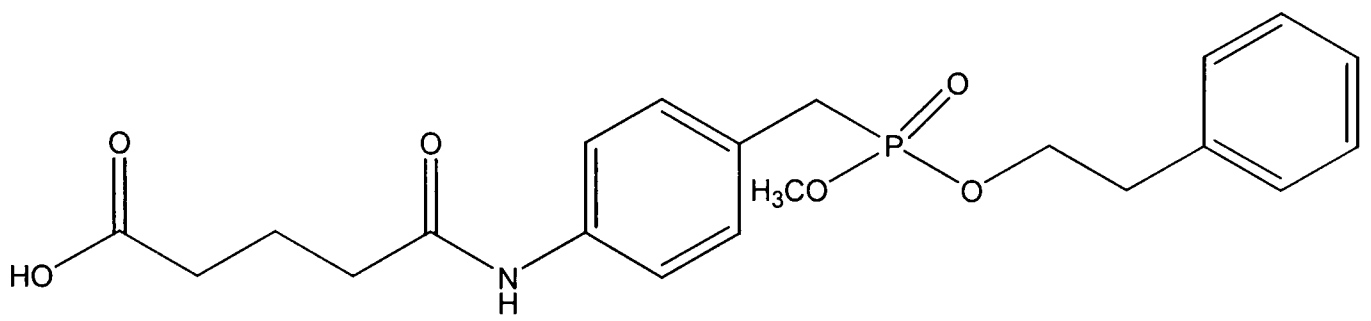
Synthesis of (4-Amino-Benzyl) Phosphonic Acid Methyl Ester Phenethyl Ester (160).



160

[4-(2,2,2-trifluoroacetylamino)-benzyl]-phosphonic acid methyl ester phenethyl ester **154** (4g, 9.97mmol) was dissolved in anhydrous MeOH (50ml) and cooled to 0°C. To the stirred solution was added sodium borohydride (3.77g, 97mmol) in portions. The solution was warmed to r.t., and stirred for 2 hours. The solvent was removed *in vacuo*, and the residue diluted with aq. NH_4OH (10%, 60ml). The aqueous phase was extracted with EtOAc (4 x 50ml). The combined organic extracts were dried over Na_2SO_4, and the solvent removed *in vacuo* to yield the product **160** as a yellow gum (2.97g, 96%); $\mathbf{R_f}$ 0.36 (SiO_2; EtOAc:MeOH; 9:1); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 2.81 (t, J=7.0, 2H, $POCH_2C\underline{H}_2Ar$), 2.92 (d, $^2J_{PH}$=21.0, 2H, $ArC\underline{H}_2P$), 3.47 (d, $^3J_{POCH}$=11.0, 3H, $OC\underline{H}_3$), 4.01-4.14 (m, 2H, $POC\underline{H}_2CH_2Ar$), 6.52 (d, J=8.0, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$), 6.93 (dd, $^4J_{PH}$=2.5, J=8.0, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.10 (d, J=7.0, 2H, *ortho*-Ar$\underline{H}$), 7.17 (t, 1H, *para*-Ar$\underline{H}$), 7.22 (t, J=7.0, 2H, *meta*-Ar$\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 31.8 (d, J_{CP}=139.0, $Ar\underline{C}H_2P$), 36.6 ($POCH_2\underline{C}H_2Ar$), 52.3 ($O\underline{C}H_3$), 66.2 ($PO\underline{C}H_2CH_2Ar$), 115.7 (C(2), C(6)), 120.7 (C(4)), 127.0 (*para*-Ar$\underline{C}$), 128.8, 129.5 (2 x *ortho*-Ar$\underline{C}$, 2 x *meta*-Ar$\underline{C}$), 130.9 (C(3), C(5)), 137.0 (1 x quaternary, Ar$\underline{C}$), 145.2 (C(1)); **^{31}P NMR** (121.4MHz, $CDCl_3$) δ_P/ppm 29.3; ν_{max} ($CHCl_3$/cm^{-1}) 650(m), 729(m), 770(m), 771(m), 1012(m), 1065(m, P-OR), 1095(m), 1271(s, P=O), 1382(m), 1459(m), 1517(w), 1623(m), 3012(m), 3619(m, -NH_2), 3665(m, -NH_2); **m/z** (APCI$^+$) 306 (100%, $[M+H]^+$); **HRMS** found 306.1259, $[M+H]^+$ ($C_{16}H_{21}NO_3P$) requires 306.1264.

Synthesis of [4-(4-Carboxy-1-Oxobutyl)-Amino-Benzyl]-Phosphonic Acid Methyl Ester Phenethyl Ester (155).

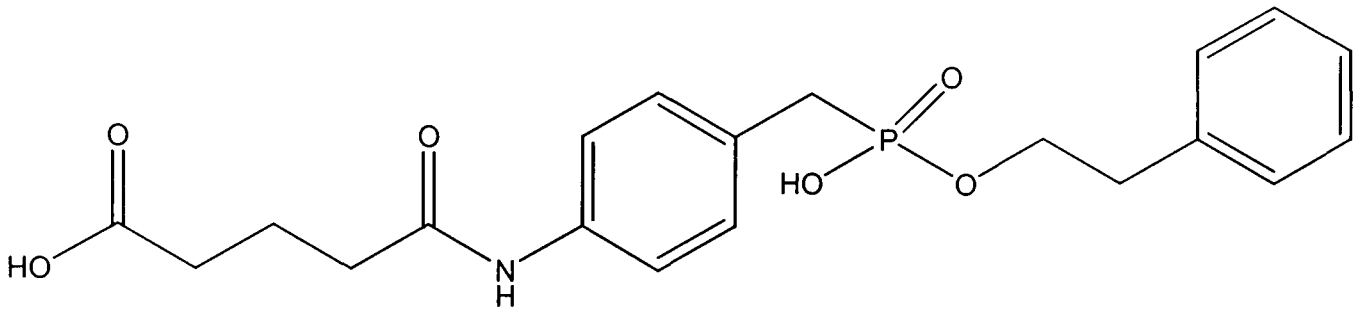


155

Triethylamine (1.73ml, 12.4mmol) and glutaric anhydride (1.35g, 12mmol) were added to a stirred solution of (4-amino-benzyl) phosphonic acid methyl ester phenethyl ester **160** (3g, 9.8mmol) in anhydrous DCM (50ml) under N_2, and the reaction stirred at reflux for 12 hours. The solution was diluted with H_2O (40ml) and adjusted to pH9 using $NaHCO_3$. The layers were separated and the aqueous layer washed with DCM (2 x 50ml). The pH of the aqueous solution was adjusted to 2 using aq. HCl (2M), and the product extracted with DCM (4 x 70ml). The organic layer was dried over Na_2SO_4, and the solvent removed *in vacuo* to afford **155** as a white gum (3.44g, 84%); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 1.99 (m, 2H, $HO_2CCH_2C\underline{H}_2CH_2C(O)NH$), 2.38, 2.41 (2t, J=7.0, 4H, $HO_2CC\underline{H}_2CH_2C\underline{H}_2C(O)NH$), 2.92 (t, J=7.0, 2H, $POCH_2C\underline{H}_2Ar$), 3.07 (d, $^2J_{PH}$=21.5, 2H, $ArC\underline{H}_2P$), 3.56 (d, $^3J_{POCH}$=11.0, 3H, $OC\underline{H}_3$), 4.18 (m, 2H, $POC\underline{H}_2CH_2Ar$), 7.12 (d, J=8.0, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.18 (d, J=7.0, 2H, *ortho*-Ar$\underline{H}$), 7.26 (t, J=7.0, 1H, *para*-Ar$\underline{H}$), 7.31 (t, J=7.0, 2H, *meta*-Ar$\underline{H}$), 7.50 (d, J=8.0, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$), 8.99 (s, 1H, C(O)N$\underline{H}$), 9.12 (s, br, $CO_2\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 21.2 ($HO_2CCH_2\underline{C}H_2CH_2C(O)NH$), 31.7, 32.4 ($HO_2C\underline{C}H_2CH_2\underline{C}H_2C(O)NH$), 33.7 ($POCH_2\underline{C}H_2Ar$), 36.9 (d, J_{CP}=139.0, $Ar\underline{C}H_2P$), 53.3 ($O\underline{C}H_3$), 67.4 ($PO\underline{C}H_2CH_2Ar$), 120.6 (C(2), C(6)), 126.3, 126.4, 127.1, 128.9 (1 x *para*-Ar$\underline{C}$, 2 x *ortho*-Ar$\underline{C}$, 2 x *meta*-Ar$\underline{C}$), 129.4 (C(3), C(5)), 130.5, 137.5, 137.9 (3 x quaternary, Ar$\underline{C}$), 172.2 ($\underline{C}$(O)NH), 176.8 ($\underline{C}O_2H$); **^{31}P NMR** (121.4MHz, $CDCl_3$) δ_P/ppm 29.5; ν_{max} ($CHCl_3$/cm^{-1}) 557(w), 650(m), 783(m), 908(m), 1018(m), 1060(m, P-OR), 1139(w), 1178(w), 1251(s, P=O), 1307(w), 1380(m), 1412(m), 1457(m), 1517(s, C(O)NH), 1603(s), 1685(s, C=O), 1703(s, C=O), 2854(m, br, CO_2H), 2954(m, br,

CO_2H), 2995(m), 3156(m), 3307(m, C(O)NH), 3429(m, C(O)NH); **m/z** ($APCI^+$) 420 (80%, $[M+H]^+$); **HRMS** found 420.1576, $[M+H]^+$ ($C_{21}H_{27}NO_6P$) requires 420.1543.

Synthesis of [4-(4-Carboxy-1-Oxobutyl)-Amino-Benzyl]-Phosphonic Acid Phenethyl Ester (146).



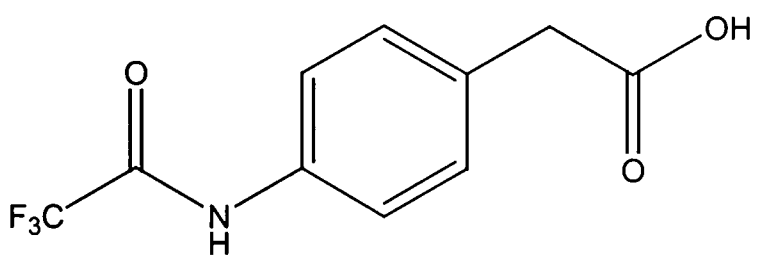
146

A suspension of **155** (2.18g, 5mmol) in *tert*-butylamine (50ml) was heated at reflux for 7 days. After this time, the excess *tert*-butylamine was removed *in vacuo*, and the residue dissolved in water and lypholised to yield a white solid. This was dissolved in MeOH (100ml) and stirred with cation exchange resin Amberlite IR-120 (plus) for 4 hours. The resin was removed by filtration, and the solvent evaporated. The residue was suspended in aq. HCl (5N, 30ml) and stirred for 24 hours. The product was collected by filtration and dried *in vacuo*, to afford **146** as a white solid (1.42g, 67%); **m.p.** 139-141°C; **^{1}H NMR** (500MHz, D_2O) δ_H/ppm 1.92 (m, 2H, $HO_2CCH_2C\underline{H}_2CH_2C(O)NH$), 2.26 (t, J=7.5, 2H, $HO_2CCH_2CH_2C\underline{H}_2C(O)NH$), 2.42 (t, J=7.5, $HO_2CC\underline{H}_2CH_2CH_2C(O)NH$), 2.85 (t, J=6.5, 2H, $POCH_2C\underline{H}_2Ar$), 2.91 (d, $^2J_{PH}$=20.5, 2H, $ArC\underline{H}_2P$), 4.10 (dt, $^3J_{POCH}$=6.5, J=6.5, 2H, $POC\underline{H}_2CH_2Ar$), 7.14 (dd, $^4J_{POCH}$=3.0, J=8.0, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.25 (d, J=7.0, 2H, *ortho*-Ar$\underline{H}$), 7.29 (t, J=7.5, 1H, *para*-Ar$\underline{H}$), 7.36 (t, J=7.5, 2H, *meta*-Ar$\underline{H}$), 7.45 (d, J=8.0, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (75.4MHz, CD_3OD) δ_C/ppm 22.5 ($HO_2CCH_2\underline{C}H_2CH_2C(O)NH$), 33.5 ($HO_2CCH_2CH_2\underline{C}H_2C(O)NH$), 34.6 ($HO_2C\underline{C}H_2CH_2CH_2C(O)NH$), 35.3 ($POCH_2\underline{C}H_2Ar$), 37.8 (d, J_{CP}=138.5, $Ar\underline{C}H_2P$), 67.8 ($PO\underline{C}H_2CH_2Ar$), 121.6, 128.0, 129.3, 129.5, 129.9, 130.6 (9 x Ar$\underline{C}$), 131.8, 139.1, 139.4 (3 x quaternary, Ar$\underline{C}$), 174.1 ($\underline{C}$(O)NH), 177.3 ($\underline{C}O_2H$); **^{31}P NMR** (121.4MHz, CD_3OD) δ_P/ppm 27.2; ν_{max} ($CHCl_3$/cm^{-1}) 494(m), 570(m), 615(m), 1009(m), 1092(s, P-OR), 1217(s), 1259(s, P=O), 1312(w), 1410(m), 1516(m, C(O)NH), 1653(s, C=O), 1698(s, C=O), 2886(m, br, CO_2H), 2977(m, br, CO_2H), 3024(m), 3621(m, C(O)NH).

3686(m, C(O)NH); **m/z** (FAB) 428 (76%, $[M+Na]^+$), 406 (42%, $[M+H]^+$); **HRMS** found 428.1239, $[M+Na]^+$ ($C_{20}H_{24}NO_6PNa$) requires 428.1225; **Anal.** Calcd. for $C_{20}H_{24}NO_6P$: C, 59.25, H, 5.96, N, 3.46, P, 7.46%, Found: C, 59.53, H, 6.01, N, 3.56, P, 8.01%.

2.1 Synthesis of the Ester Substrate.

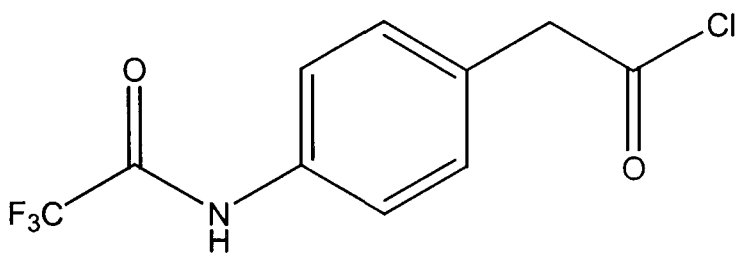
Synthesis of [4-(2,2,2-Trifluoro-Acetylamino)-Phenyl]-Acetic Acid (164)[161].



164

4-Aminophenylacetic Acid **163** (10g, 0.066mol) was dissolved in acetonitrile (90ml) and H_2O (10ml) and cooled to 0°C. Na_2CO_3 (13.87g, 0.132mol) was added, followed by the drop-wise addition of trifluoroacetic anhydride (28.03ml, 0.198mol). The reaction was allowed to warm to r.t., and then stirred vigorously for 1 hour. The acetonitrile was removed *in vacuo*, and the residue diluted with EtOAc (100ml) and H_2O (100ml). The pH was adjusted to 2 using conc. HCl. The layers were separated, and the organic phase washed with aq. HCl (2M, 2 x 50ml), aq. NaCl (sat., 1 x 50ml), dried over Na_2SO_4, and concentrated until a white precipitate appeared. The solid was collected by filtration to afford [4-(2,2,2-trifluoro-acetylamino)-phenyl]-acetic acid **164** as a white crystalline solid (9.4g, 58%); **m.p.** 183-185°C (lit. 175°C[161]); **^{1}H NMR** (300MHz, CD_3OD) δ_H/ppm 3.60 (s, 2H, Ar$\underline{C}H_2CO_2H$), 7.3 (d, *J*=8.5, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.56 (d, *J*=8.5, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (75.4MHz, CD_3OD) δ_C/ppm 41.7 (Ar$\underline{C}H_2CO_2H$), 116.0 (1 x quaternary, $\underline{C}F_3C(O)NH$), 122.6, 122.7 ($\underline{C}$(2), $\underline{C}$(6)), 131.4, 131.5 ($\underline{C}$(3), $\underline{C}$(5)), 134.2, 136.9 (2 x quaternary, Ar$\underline{C}$), 157.4 ($CF_3\underline{C}(O)NH$), 175.8 ($\underline{C}O_2H$); ν_{max} (KBr/cm^{-1}) 671(m), 712(w), 794(w), 835(m), 1148(m, C-F), 1176(s, C-F), 1250(m), 1291(m), 1346(m), 1421(s), 1543(s), 1598(m), 1700(s, C=O), 3290(s, C(O)NH); **m/z** (FAB) 258 (100%, $[M+H]^+$); **HRMS** found 248.0540, $[M+H]^+$ ($C_{10}H_9NO_3F_3$) requires 248.0535.

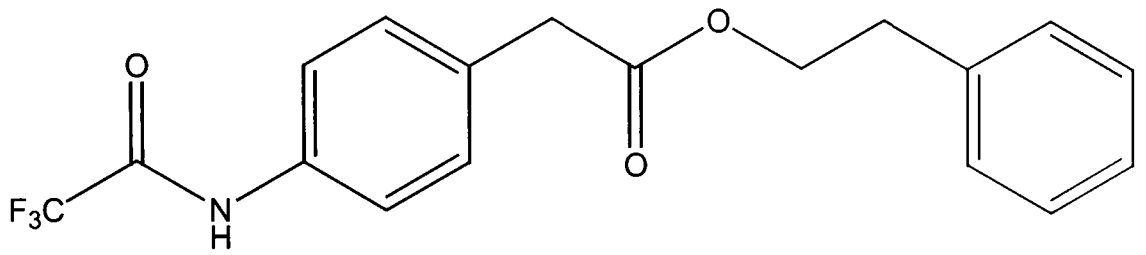
Synthesis of [4-(2,2,2-Trifluoro-Acetylamino)-Phenyl]-Acetyl Chloride (165)[162].



165

To a stirred suspension of [4-(2,2,2-trifluoro-acetylamino)-phenyl]-acetic acid **164** (3g, 0.012mol) in anhydrous DCM (80ml) under N_2 was added oxalyl chloride (1.38ml, 0.016mol) and DMF (40µl, trace). The suspension was stirred for 3 hours, after which time it had become a homogenous solution. The solvent and excess reagents were removed *in vacuo* to afford [4-(2,2,2-trifluoro-acetylamino)-phenyl]-acetyl chloride **165** as an off-white solid (3.15g, 99%); **^{1}H NMR** (400MHz, CD_3CN) δ_H/ppm 4.28 (s, 2H, Ar$\underline{C}H_2$COCl), 7.33 (d, *J*=8.5, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.62 (d, *J*=8.5, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (100.6MHz, CD_3CN) δ_C/ppm 52.7 (Ar$\underline{C}H_2$COCl), 118.3 ($\underline{C}$(2), $\underline{C}$(6)), 122.3 (1 x quaternary, $\underline{C}F_3C$(O)NH), 130.8 (1 x quaternary, $\underline{C}$(4)), 131.4 ($\underline{C}$(3), $\underline{C}$(5)), 136.7 (1 x quaternary, $\underline{C}$(1)), 157.3 ($CF_3\underline{C}$(O)NH), 173.5 ($\underline{C}$OCl); ν_{max} (KBr/cm^{-1}) 509(m), 561(m), 599(s), 636(m), 710(s), 800(m), 1002(s), 1151(s), 1203(s), 1247(m), 1285(m), 1404(m), 1502(m), 1554(m, C(O)NH), 1614(m), 1704(s, C=O), 1793(s, C=O), 3102(m), 3147(m, C(O)NH), 3214(m, C(O)NH), 3297(m, C(O)NH); **m/z** (FAB) 265, 267 (30%, $[M]^+$).

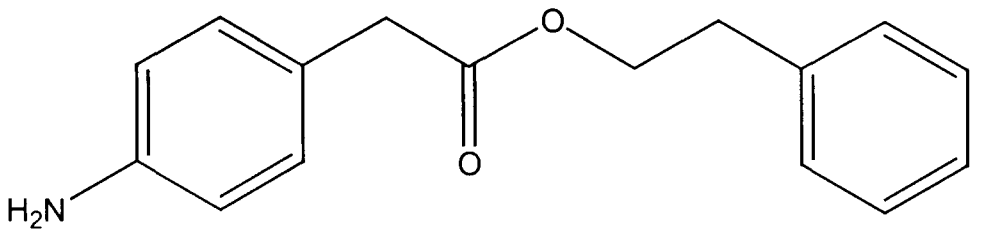
Synthesis of [4-(2,2,2-Trifluoro-Acetylamino)-Phenyl]-Acetic Acid Phenethyl Ester (166).



166

A solution of [4-(2,2,2-trifluoro-acetylamino)-phenyl]-acetyl chloride **165** (513mg, 1.93mmol) in anhydrous THF (10ml) was added slowly to a stirred solution of 2-phenylethanol (346µl, 2.9mmol) and triethylamine (807µl, 5.8mmol) in anhydrous THF (20ml) under N_2. The reaction was stirred for 12 hours. Concentration *in vacuo*, gave a white solid which was dissolved in EtOAc (100ml), then washed with aq. HCl (0.5M, 2 x 50ml), aq. $NaHCO_3$ (sat., 2 x 50ml), aq. NaCl (sat., 1 x 50ml), and dried over Na_2SO_4. The solvent was removed *in vacuo* to yield the crude product as an orange oil. Purification by flash chromatography (SiO_2; DCM) afforded **166** as a white solid (430mg, 63%); **m.p.** 96-98°C; $\mathbf{R_f}$ 0.37 (SiO_2; DCM); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 2.81 (t, *J*=7.0, 2H, $CO_2CH_2C\underline{H}_2Ar$), 3.48 (s, 2H, $ArC\underline{H}_2CO_2R$), 4.20 (t, *J*=7.0, 2H, $CO_2C\underline{H}_2CH_2Ar$), 7.12 (d, *J*=8.5, 2H, C(3)<u>H</u>, C(5)<u>H</u>), 7.15-7.22 (m, 5H, 5 x Ar<u>H</u>), 7.37 (d, *J*=8.5, 2H, C(2)<u>H</u>, C(6)<u>H</u>), 8.09 (s, 1H, $CF_3C(O)N\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 35.4 ($CO_2CH_2\underline{C}H_2Ar$), 41.2 ($Ar\underline{C}H_2CO_2R$), 65.9 ($CO_2\underline{C}H_2CH_2Ar$), 120.9, 121.0 (<u>C</u>(2), <u>C</u>(6)), 124.5 (1 x quaternary, $\underline{C}F_3C(O)NH$), 127.0 (*para*-Ar<u>C</u>), 128.8, 128.9, 129.3, 129.4 (2 x *ortho* Ar<u>C</u>, 2 x *meta* Ar<u>C</u>), 130.6, 130.7 (<u>C</u>(3), <u>C</u>(5)), 132.5, 134.5, 137.9 (3 x quaternary, Ar<u>C</u>) 157.4 ($CF_3\underline{C}(O)NH$), 171.7 ($\underline{C}O_2H$); ν_{max} ($CHCl_3$/cm^{-1}) 650(m), 708(s), 752(s), 767(m), 1098(m, C-OR), 1250(m, C-OR), 1278(m), 1287(m), 1342(w), 1380(m), 1465(m), 1543(m), 1605(m), 1744(s, C=O), 1792(s, C=O), 2900(m), 2959(m), 2978(m), 3149(m), 3422(m, C(O)NH), 3617(m, C(O)NH); **m/z** (FAB) 352 (40%, $[M+H]^+$), 374 (70%, $[M+Na]^+$); **HRMS** found 374.0969, $[M+Na]^+$ ($C_{18}H_{16}F_3NO_3Na$) requires 374.0980.

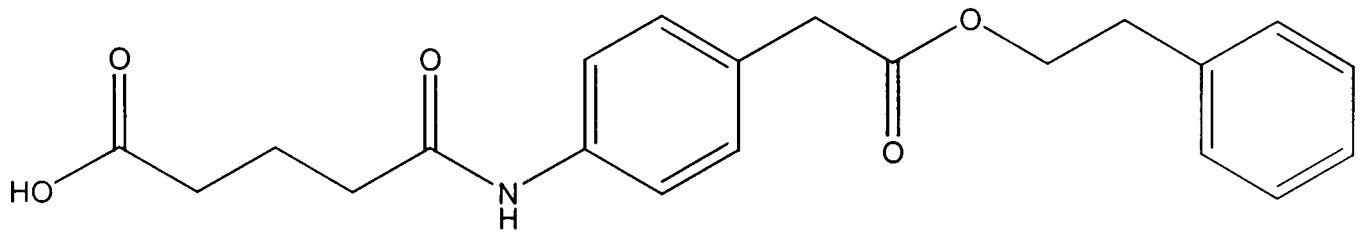
Synthesis of 4-Amino-Phenyl Acetic Acid Phenethyl Ester (167).



167

Sodium borohydride (564mg, 15mmol) was added to a stirred solution of **166** (2.62g, 7.5mmol) in anhydrous MeOH (50ml), and the solution stirred at r.t. for 14 hours. The solvent was removed *in vacuo*, and the residue diluted with EtOAc (200ml), and aq. NH_4OH (10%, 200ml). The layers were separated and the aqueous layer washed with EtOAc (2 x 60ml). The organic layers were combined, and the product extracted into aq. HCl (2M, 3 x 75ml). The pH was adjusted to 9 using $NaHCO_3$, and the product extracted into EtOAc (3 x 75ml), dried over Na_2SO_4, and the solvent removed *in vacuo* to afford a yellow oil. Purification by flash chromatography (SiO_2; DCM:EtOAc; 11:1) afforded **167** as a yellow oil (0.75g, 40%); $\mathbf{R_f}$ 0.4 (SiO_2; DCM:EtOAc; 11:1); **^{1}H NMR** (300MHz, $CDCl_3$) δ_H/ppm 2.92 (t, *J*=7.0, 2H, $CO_2CH_2C\underline{H}_2Ar$), 3.49 (s, 2H, $ArC\underline{H}_2CO_2R$), 4.31 (t, *J*=7.0, 2H, $CO_2C\underline{H}_2CH_2Ar$), 6.60, (d, *J*=8.5, 2H, C(2)H, C(6)H), 7.03 (d, *J*=8.5, 2H, C(3)H, C(5)H), 7.08 (1H, m, *para*-ArH), 7.18 (d, *J*=7.5, 2 x *ortho*-ArH), 7.28 (2H, m, 2 x *meta*-ArH); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 34.7 ($CO_2CH_2\underline{C}H_2Ar$), 40.2 ($Ar\underline{C}H_2CO_2R$), 64.9 ($CO_2\underline{C}H_2CH_2Ar$), 114.9 (C(2), C(6)), 123.3 (1 x quaternary, C(4)), 126.2 (*para*-ArC), 128.2, 128.7 (2 x *ortho*-ArC, 2 x *meta*-ArC), 129.8, 129.9 (C(3), C(5)) 137.6 (1 x quaternary ArC), 145.4 (1 x quaternary, C(1)), 171.8 ($\underline{C}O_2R$); ν_{max} (thin film/cm^{-1}) 499(w), 519(w), 700(m), 749(m), 833(m), 1003(m), 1150(s, C-OR), 1254(m, C-OR), 1435(w), 1453(w), 1517(m), 1626(m), 1725(s, C=O), 2955(m), 3028(w), 3060(w), 3350(w), 3371(s, $-NH_2$), 3453(s, $-NH_2$); **m/z** ($APCI^+$) 256.0 (100%, $[M+H]^+$).

Synthesis of Benzene Acetic Acid 4-[(4-Carboxy-1-Oxobutyl)amino]-Phenethyl Ester (145, Ester Substrate).



145

Glutaric anhydride (0.376g, 3.3mmol) was added to a stirred solution of triethylamine (483µl, 3.5mmol) and 4-amino-phenyl acetic acid phenethyl ester **167** (0.7g, 2.8mmol) in anhydrous DCM (25ml) under N_2. The solution was heated at reflux for 16 hours. After cooling the solvent was removed *in vacuo*, and the residue dissolved in EtOAc (50ml). The solution was washed with aq. HCl (2M, 1 x 50ml), and aq. NaCl (sat., 1 x 50ml), dried over Na_2SO_4, and the solvent removed *in vacuo* to afford a beige solid as the crude product. The product was recrystallised from H_2O/CH_3CN to yield **145** as a white solid (0.99g, 99%); **m.p.** 105-107°C; $\mathbf{R_f}$ 0.15 (SiO_2; DCM:EtOAc; 11:1); **^{1}H NMR** (400MHz, $CDCl_3$) δ_H/ppm 2.03 (m, 2H, $HO_2CCH_2C\underline{H}_2CH_2C(O)NH$), 2.43 (t *J*=7.5, 2H, $HO_2CCH_2CH_2C\underline{H}_2C(O)NH$), 2.47 (t, *J*=7.5, 2H, $HO_2C\underline{C}H_2CH_2CH_2C(O)NH$), 2.95 (t, *J*=7.0, 2H, $CO_2CH_2C\underline{H}_2Ar$) 3.59 (s, 2H, $ArC\underline{H}_2CO_2R$), 4.33 (t, *J*=7.0, 2H, $CO_2C\underline{H}_2CH_2Ar$), 7.16 (d, *J*=8.5, C(2)<u>H</u>, C(6)<u>H</u>), 7.18 (d, *J*=7.0, 2H, 2 x *ortho*-Ar<u>H</u>), 7.25 (1H, m, *para*-Ar<u>H</u>), 7.28 (2H, m, 2 x *meta*-Ar<u>H</u>), 7.46 (d, *J*=8.5, C(3)<u>H</u>, C(5)<u>H</u>), 8.00 (s, 1H, C(O)N<u>H</u>); **^{13}C NMR** (100.6MHz, $CDCl_3$) δ_C/ppm 20.5 ($HO_2CCH_2\underline{C}H_2CH_2C(O)NH$), 32.9 ($HO_2CCH_2CH_2\underline{C}H_2C(O)NH$), 34.9 ($CO_2CH_2\underline{C}H_2Ar$), 36.0 ($HO_2\underline{C}CH_2CH_2CH_2C(O)NH$), 40.7 ($Ar\underline{C}H_2CO_2R$), 65.4 ($CO_2\underline{C}H_2CH_2Ar$), 120.1 (<u>C</u>(2), <u>C</u>(6)), 126.5 (*para*-Ar<u>C</u>), 128.5, 128.9 (2 x *ortho*-Ar<u>C</u>, 2 x *meta*-Ar<u>C</u>), 129.8 (<u>C</u>(3), <u>C</u>(5)), 130.6, 136.8, 137.5 (3 x quaternary, Ar<u>C</u>), 171.1 ($\underline{C}O_2R$), 171.8 (<u>C</u>(O)NH), 177.9 ($\underline{C}O_2H$); ν_{max} (nujol/cm^{-1}) 794(w), 941(w), 1166(m), 1224(m), 1313(w), 1348(w), 1377(s), 1411(m), 1460(s, nujol), 1533(s, C(O)NH), 1598(m), 1654(s, C=O), 1691(s, C=O), 1730(s, C=O), 2677(m, br, CO_2H), 2870(m, br, CO_2H), 2852(s, nujol), 2923(s, nujol), 3319(m, C(O)NH); **m/z** (FAB) 392 (75%, $[M+Na]^+$), 370 (52%, $[M+H]^+$), 286 (50%), 255 (33%); **HRMS** found 370.1647,

$[M+H]^+$ ($C_{21}H_{25}NO_5$) requires 370.1654; **Anal.** Calcd. for $C_{21}H_{24}NO_5$: C, 68.28; H, 6.28, N 3.79%, Found: C, 68.19; H, 6.22; N, 3.72%.

3. Binding Studies - Diffusion Measurements by NMR, and Molecular Modelling.

3.0 Measurement of Diffusion Coefficients of Peptides and TSA (146) Using Pulsed Field Gradient (NMR) Techniques.

The maximum field strength of the gradient accessory was determined to be $50Gcm^{-1}$ according to standard procedure[163].

3mM solutions of each dipeptide (**168-176,** Chapter 2.4) in D_2O were prepared. The pD was adjusted to 7 using NaOD and DCl. A 30mM solution of TSA **146** in D_2O at pD7 was prepared in the same way. Sample mixtures containing one peptide and the TSA **146** were also made up so that the concentration of the peptide was 3mM, and TSA 30mM, and the pD adjusted to 7 using the method above.

Diffusion coefficients of the peptides and the TSA, both in the pure samples, and the mixtures were measured by PFG-NMR spectroscopy using the bipolar pulse pair longnitudinal encode-decode (BPP-LED) sequence[125]. This pulse sequence contains a delay T_e (11.2ms in our experiments) at the end of the experiment which avoids special artefacts arising from residual eddy currents.

In these experiments, the attenuation of the NMR resonance depends on gradient areas as shown in Equation 2:

$$I = I_o\exp[- D(\Delta - \delta/3 - \tau/2)\gamma^2 g^2 \delta^2] \qquad \textbf{Equation 2}$$

Where I_o is the intensity of the resonance in the NMR spectrum in the absence of gradient pulses, γ is the gyromagnetic ratio, Δ is the diffusion delay time, g and δ are the amplitude and duration of the bipolar gradient pulse pair, D is the diffusion coefficient, and τ is the delay between the bipolar gradient pulse pair.

In each experiment, a data set of 32 separate BPP-LED spectra was acquired as a function of gradient amplitude which ranged from 5-95% (100% $=50Gcm^{-1}$). The spacing of the gradient amplitudes was incremented according to a "squared scale" such that spacing between the squared gradient is equal. The following variables were kept

constant: δ=1.2ms, τ=01ms, Δ=200ms. All of the diffusion measurements were performed at 298K. In all experiments, the HOD resonance was suppressed by application of a selective saturation pulse to improve the reliability of the measurement of α-proton intensities and to increase the value of the receiver gain which could be used in the acquisition of the FID.

Data processing was performed on standard T_1/T_2 software under XwinNMR which utilises a standard simplex algorithm to fit data to Equation 2. The results are shown in Results and Discussion. The error value for the experiments is $< \pm 0.1 \times 10^{-10} m^2/s$.

As a control, we always checked the diffusion coefficient of H_2O which should have a value of 2.3×10^{-9} m^2/s at 298K[156,157]. In our experiments a value between 2.2 and 2.4 x $10^{-9} m^2/s$ was always observed, even when water suppression was performed on the residual water peak.

3.1 Determination of the Binding Constant Ka for the complex between TSA 146 and Arginine-Arginine.

11 Samples were prepared all with a 1mM concentration of arginine-arginine **168,** and with a TSA **146** concentration varying from 0-40mM. The diffusion coefficients of the TSA **146** and the peptide **168** were measured in all the samples using the method described above (3.0). The Associate 1.6 programme, which employs the non-linear least squares Levenberg-Marquardt method to fit parameters to equilibrium complexation models, was used to determine the binding constant K_a[129,130]. The results are shown in Chapter 2.4

3.2 Determination of Internuclear Distances in the TSA Using 2D-NOESY Experiments.

A 30mM solution of TSA **146** in D_2O was prepared. The pD was adjusted to 7 using DCl and NaOD. 2D-NOESY experiments were carried out at 298K using 256 experiments in F1 Dimension, and mixing delays (τ_m) varying from 100-600ms. From these spectra integral volume intensities of all of the cross peaks were measured using standard XwinNMR routines. Based on the fact that the distance for *ortho* protons of a

phenyl group is normally 2.8Å (r_{ref}), Equation 8 was used to calculate the internuclear distance (r_{is}) between two nuclei a and b:

$$r_{ab} = r_{ref}\left[\frac{\text{intensity}_{ab}}{\text{intensity}_{ref}}\right]^{-\frac{1}{6}} \qquad \textbf{Equation 8}$$

where intensity_{ab} is the cross peak intensity of the two spins to be measured at time t, intensity_{ref} is the reference cross peak intensity, r_{ref} is the reference distance and r_{ab} is the unknown distance between nuclei a and b. The results are as shown in Chapter 2.4, Table 3.

The accuracy of this method is reasonably reliable because it relies on the sixth root of NOE intensities, thus large errors in intensity measurements have a small effect on the derived internuclear seperation[136].

3.3 Molecular Modelling.

Arginine-arginine **168** was built using the Biopolymer module of Sybyl with extended conformations. [4-(4-Carboxy-1-oxobutyl)-amino-benzyl]-phosphonic acid phenethyl ester (TSA) **146** was sketched in Sybyl[131] and converted into 3D using Concord. Both molecules were minimised using the Tripos force field with Gasteiger-Huckel charges and Powell minimiser to an RMS gradient of 0.005kcal/Å, Non-bonded cut-offs were applied at 8Å and a distance dependant dielectric of 80 was used to crudely approximate the electrostatic screening due to solvation. Both molecules were then subjected to a conformational search using the Random search facility of Sybyl for 5000 iterations using the following parameters: Energy cut-off 500kcal/mol; Convergence threshold 0.05; RMS threshold 0.75 Å; Minimum hits 200. Each conformation was minimised for 100 steps or to gradient less than 0.05 kcal/ Å using the parameters described above.

The geometrical constraints measured by NOESY NMR were applied to the database of conformations generated for the TSA. Four conformations fitted the constraints, and of these two were discarded due to distortions of one of the phenyl rings. The distance between the carboxylate and phosphonate groups on the remaining two conformers was measured and found to be 12-14 Å.

For the Arg-Arg **168** to complex as a bidentate ligand to the TSA **146** the distance between either the two guanidine groups, or one guanidine group and the terminal

amino group had to be similar to this distance. This constraint was applied to the database of conformations generated for Arg-Arg and gave ten conformations to dock against the TSA.

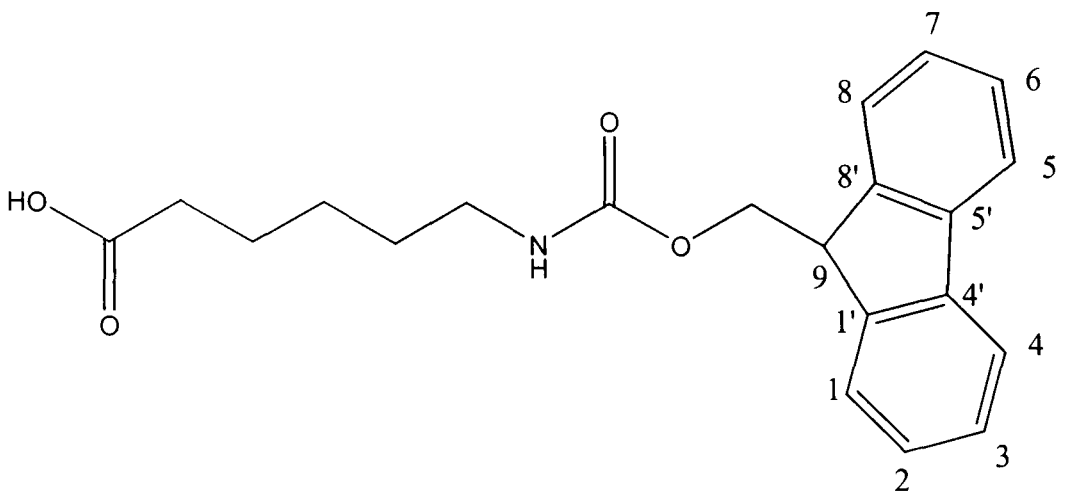
All combinations of the above conformations of both molecules were fitted together using the DOCK command until the electrostatic interactions were maximised, and the steric energy minimised. The complex was then minimised using the parameters above. The energies of both the complex (E_{total}), and the individual components were measured, and used to calculate the interaction energy of the complex (E_i) using Equation 9:

$$E_i=E_{total}-(E_{pep} + E_{TSA}) \qquad \textbf{Equation 9}$$

Where E_{pep} is the energy of the peptide, and E_{TSA} is the energy of the TSA. The results are shown in Chapter 2.4, Tables 4-6.

4. Synthesis of 'Tripeptide Binding Unit'.

4.0 Synthesis of 6-(*9*H-Fluoren-9-ylmethoxycarbonylamino)-Hexanoic Acid (182).

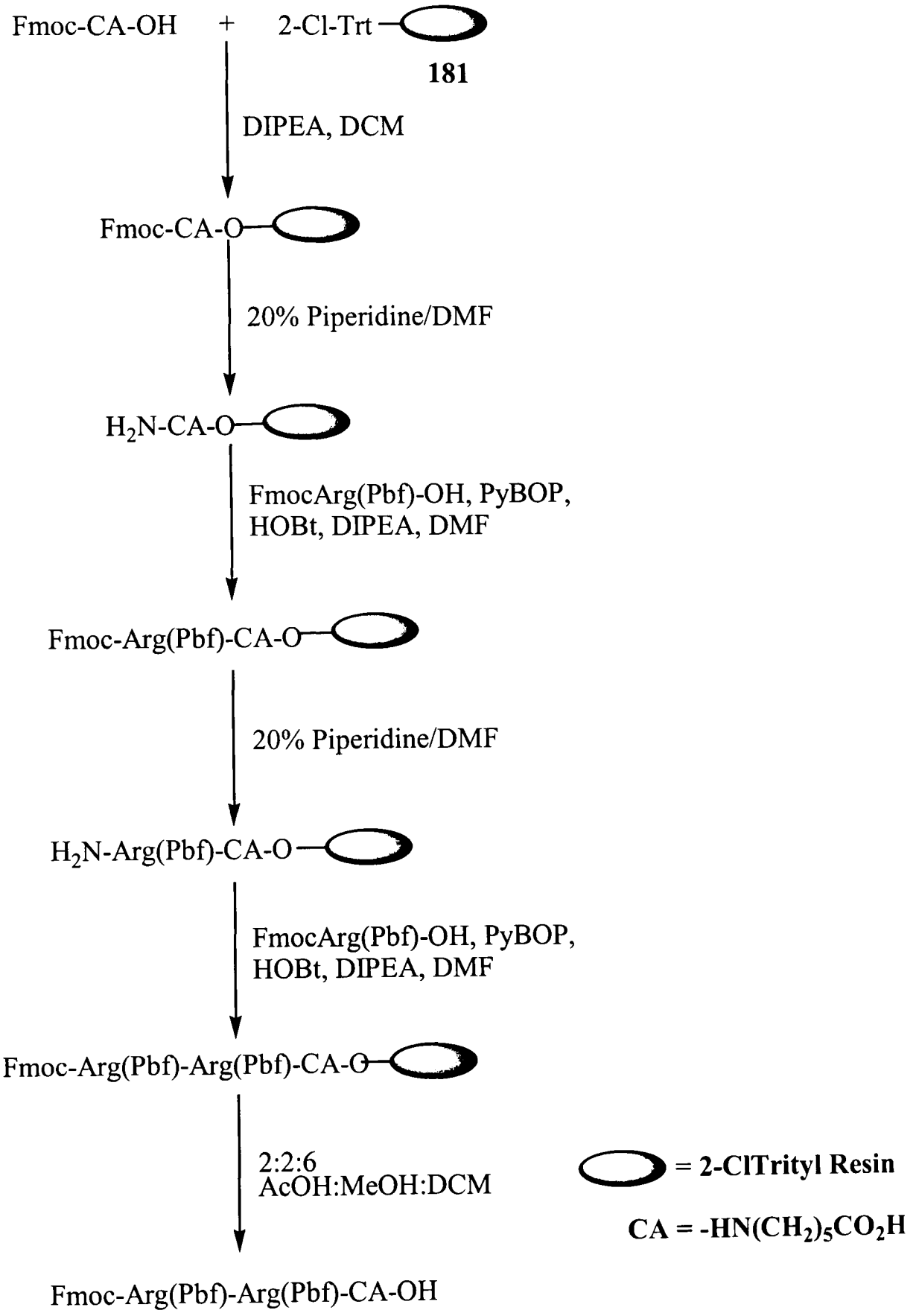


182

A stirred solution of 6-aminocaproic acid **179** (3g, 23mmol) in 10% aq. Na_2CO_3 solution (6.09g, 57mmol in 60ml H_2O) was cooled to 0°C. To the cooled solution was added 9-fluorenylmethyl chloroformate (5.92g, 23mmol) in dioxane (30ml). The solution was stirred for 10 mins., after which time a white precipitate had formed, and then allowed to warm to r.t. and stirred for 4 hours. The mixture was diluted with H_2O (500ml), and washed with Et_2O (2 x 100ml). The aqueous layer was adjusted to pH 1 using conc. HCl, and a white precipitate formed, which was extracted into EtOAc (3 x 200ml). The organic layers were combined and dried over Na_2SO_4, and the solvent removed *in vacuo* to yield the crude product as a white solid. Purification by flash chromatography (SiO_2; EtOAc) yielded 6-(*9*H-fluoren-9-ylmethoxycarbonylamino)-hexanoic acid **182** as a white solid (6.98g, 86%); **m.p.** 118-120°C; $\mathbf{R_f}$ 0.58 (SiO_2; EtOAc); **^{1}H NMR** (500MHz, $CDCl_3$) δ_H/ppm 1.37 (m, 2H, $HO_2CCH_2CH_2C\underline{H}_2CH_2$), 1.51 (m, 2H, $HO_2CCH_2CH_2CH_2C\underline{H}_2$), 1.67 (m, 2H, $HO_2CCH_2C\underline{H}_2CH_2CH_2$), 2.34 (t, J=7.0, $C\underline{H}_2CO_2H$), 3.04 (br, s, minor rotamer, $C\underline{H}_2NHC(O)$), 3.17 (q, J=6.5, major rotamer, $C\underline{H}_2NHC(O)$), 4.19 (t, J=6.5, 1H, C(9)$\underline{H}$), 4.38 (d, J=6.5, major rotamer, $C(9)HC\underline{H}_2OC(O)$), 4.45 (br, s, minor rotamer, $C(9)HC\underline{H}_2OC(O)$), 4.77 (br, s, 1H, $CH_2N\underline{H}C(O)$), 7.29 (t, J=7.5, 2H, C(2)$\underline{H}$, C(7)$\underline{H}$), 7.38 (t, J=7.5, 2H, C(3)$\underline{H}$, C(6)$\underline{H}$), 7.57 (d, J=7.5, 2H, C(1)$\underline{H}$, C(8)$\underline{H}$), 7.70 (d, J=7.5, 2H, C(4)$\underline{H}$, C(5)$\underline{H}$); **^{13}C NMR** (125.7MHz, $CDCl_3$) δ_C/ppm 24.2, 26.1, 29.5 (3 x CH_2, $HO_2CCH_2\underline{C}H_2\underline{C}H_2\underline{C}H_2$), 33.8

($\underline{C}H_2NHC(O)$), 40.7 ($HO_2C\underline{C}H_2CH_2CH_2CH_2$), 47.2 ($\underline{C}(9)H$), 66.9 ($C(9)H\underline{C}H_2OC(O)$), 119.9, 125.0, 127.0, 127.6 ($\underline{C}(1)$, $\underline{C}(2)$, $\underline{C}(3)$, $\underline{C}(4)$, $\underline{C}(5)$, $\underline{C}(6)$, $\underline{C}(7)$, $\underline{C}(8)$), 141.3 ($\underline{C}(1')$, $\underline{C}(8')$), 143.9 ($\underline{C}(4')$, $\underline{C}(5')$), 156.4 ($NH\underline{C}(O)O$), 179.0 ($\underline{C}O_2H$); ν_{max} (nujol/cm^{-1}) 732(s), 779(m), 949(m), 1002(m), 1031(m), 1078(w), 1102(m), 1131(m), 1249(s), 1260(m), 1296(m), 1343(m), 1372(m), 1454(s, nujol), 1525(s, C=O), 1684(s, C=O), 1707(s, C=O), 2660(m, CO_2H, str), 2719(m, CO_2H, str), 2848(s, CH_2, nujol), 2919(s, CH_2, nujol), 3342(s, OC(O)NH); **m/z** (FAB) 376 (26%, $[M+Na]^+$), 354 (16%, $[M+H]^+$), 179 (100%); **HRMS** found 376.1525, $[M+Na]^+$ ($C_{21}H_{23}NO_4Na$) requires 376.1517; **Anal.** Calcd. for $C_{21}H_{23}NO_4$: C, 71.37; H, 6.56; N, 3.96%, Found: C, 71.39; H, 6.57; N, 3.88%.

4.1 Synthesis of 6-(Fmoc-Arg(Pbf)-Arg(Pbf))-Amino Hexanoic Acid (186).



Coupling of 6-(9H-Fluoren-9-ylmethoxycarbonylamino)-Hexanoic Acid to 2-Chlorotrityl Resin.

DIPEA (1.81ml, 10.4mmol) was added to a solution of 6-(9H-Fluoren-9-ylmethoxycarbonylamino)-hexanoic acid **182** (0.91g, 2.6mmol) in anhydrous DCM (20ml) under N_2, and the solution stirred for 10mins. The solution was added to 2-chlorotrityl resin **181** (substitution 1.08mmol/g) under N_2, and the mixture agitated by sonication at r.t. for 3 hours. After this time the resin was filtered, and washed with MeOH (3 x 50ml), MeOH:DCM:DIPEA (2:17:1, 3 x 50ml), DCM (3 x 50ml), DMF (3 x 50ml) and DCM (3 x 50ml), and dried under vacuum (2.91g).

Test For Level of Fmoc Substitution.

Two Samples of dry resin (0.5-2mg) were weighed into 10mm silica UV cells and dissolved in 20% Piperidine/DMF (3ml). The samples were agitated for 10mins. using a Pasteur pipette. The UV absorbance at 290nm for each sample was recorded, and the loading calculated using Equation 11:

Loading (mmol/g) = (Abs/mass of sample) x 1.75 **Equation 11**

The loading was found to be 0.88mmol/g. This value was used to calculate no. of mols. of reagents to be used in subsequent couplings.

Fmoc Deprotection.

The resin (2.91g) was suspended in 20% piperidine/DMF (25ml), and sonicated at r.t. for 3 hours. The resin was filtered, and washed with DMF (5 x 60ml), and DCM (5 x 50ml), and dried under vacuum (1.902g).

1st Coupling of Fmoc-Arg(Pbf)-OH to Resin.

Fmoc-Arg(Pbf)-OH (3.25g, 5mmol), PyBOP (2.61g, 5mmol), and HOBt (677mg, 5mmol) were dissolved in anhydrous DMF (40ml) under N_2. DIPEA (1.75ml, 10mmol) was added, and the solution added immediately to the resin (1.9g, 1.67mmol). The suspension was agitated by sonication at r.t. for 4 hours. After this time the Kaiser test indicated complete reaction of the resin bound free amine. The resin was filtered, and washed with DMF (4 x 100ml), DCM (3 x 100ml), DMF (2 x 50ml), and DCM (2 x

100ml). The resin was dried under vacuum (2.734g). The test for level of Fmoc substitution was performed using the above method, and found to be 0.6mmol/g.

Fmoc Deprotection.

The resin (2.7g) was suspended in 20% piperidine/DMF (25ml), and sonicated at r.t. for 3 hours. The resin was filtered, and washed with DMF (4 x 75ml), and DCM (4 x 75ml), and dried under vacuum (2.00g)

2nd Coupling of Fmoc-Arg(Pbf)-OH to Resin.

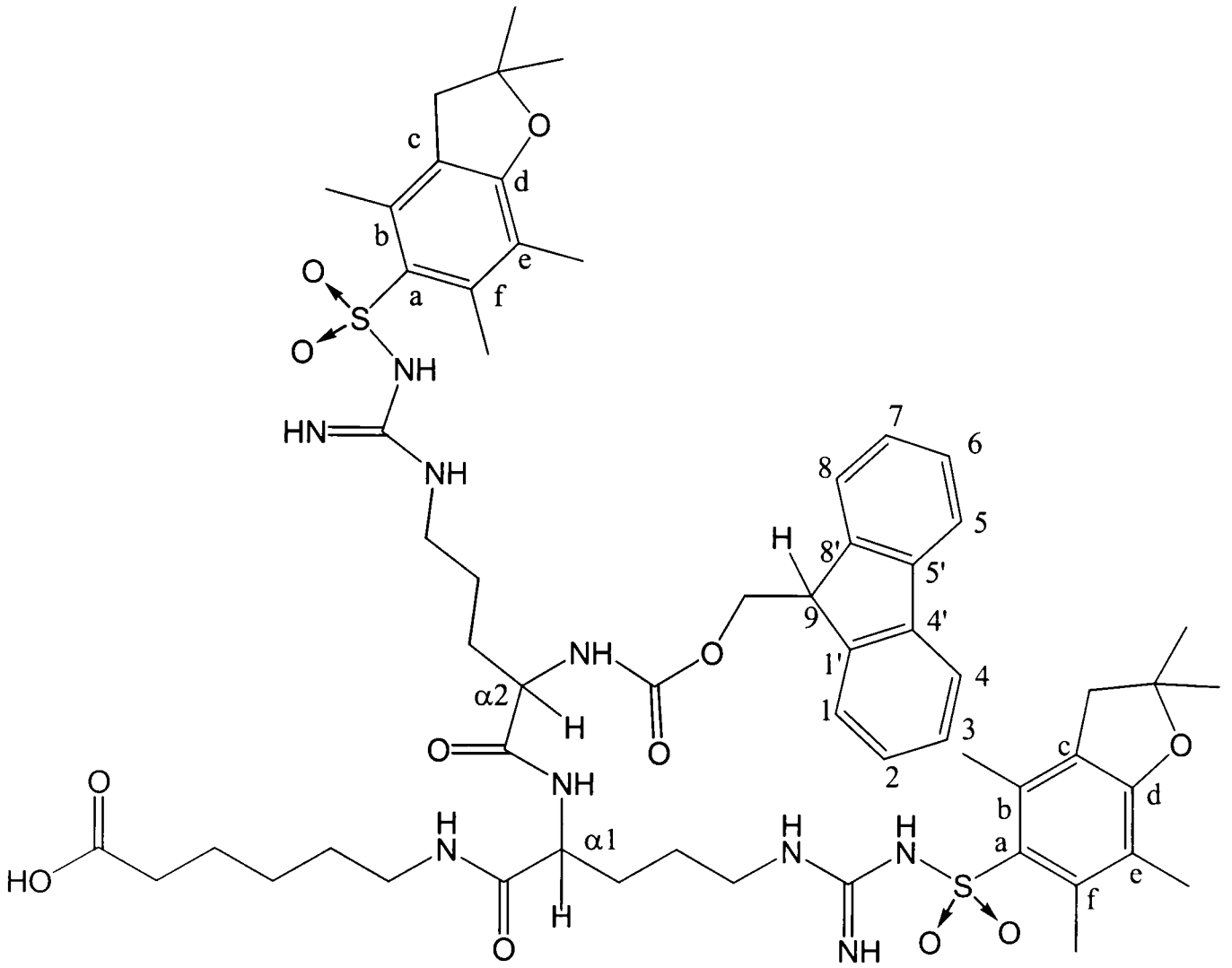
Fmoc-Arg(Pbf)-OH (1.74g, 2.7mmol), PyBOP (1.4g, 2.7mmol), and HOBt (362mg, 2.7mmol) were dissolved in anhydrous DMF (30ml) under N_2. DIPEA (0.95ml, 5.4mmol) was added and the solution added immediately to the resin (1g, 0.7mmol) under N_2. The suspension was sonicated at r.t. for 4 hours. After this time the Kaiser test indicated complete reaction of the resin bound amine. The resin was filtered, and washed with DMF (4 x 75ml), DCM (3 x 75ml), DMF (2 x 50ml), and DCM (2 x 75ml). The resin was dried under vacuum (1.4g). The test for level of Fmoc substitution was performed using the above method, and found to be 0.3mmol/g.

Cleavage of Peptide from Resin.

The resin (1.4g) was sonicated in a solution of AcOH:trifluoroethanol:DCM (2:2:6, 20ml) for 3 hours. The resin was filtered, and washed with the above solution (1 x 50ml), DCM (1 x 50ml), MeOH (1 x 30ml), and DCM (1 x 50ml). The solution was diluted with hexane (250ml), and the solvent removed *in vacuo* to afford a cream coloured solid. Flash chromatography (SiO_2; EtOAc:MeOH; 85:15) yielded **186** as a cream solid (0.58g).

Analytical data for 6- (Fmoc-Arg(Pbf)-Arg(Pbf)-)aminohexanoic acid 186 (Fmoc-Arg(Pbf)Arg(Pbf)CA-OH).

Numbering and lettering is for analytical purposes only.



186

m.p. 135-137°C; $\mathbf{R_f}$ 0.55 (SiO_2; EtOAc: MeOH; 80:20); **^{1}H NMR** (500MHz, CD_3OD) δ_H/ppm 1.36-1.56 (m, 14H, $HO_2CCH_2C\underline{H}_2C\underline{H}_2C\underline{H}_2$, 2 x ($C\underline{H}_2C\underline{H}_2CH_2NH(=NH)$)), 1.41 (s, 12H, 2 x(-$CH_2C(C\underline{H}_3)_2O$-)), 2.04, 2.05 (2s, 2 rotamers, 6H, 2 x (C(e)$C\underline{H}_3$)), 2.22 (t, *J*=7.0, 2H, $HO_2CC\underline{H}_2$), 2.49, 2.50 (2s, 2 rotamers, 6H, 2 x (C(f)$C\underline{H}_3$)), 2.55, 2.56 (2s, 2 rotamers, 6H, 2 x (C(b)$C\underline{H}_3$)), 2.94 (d, *J*=4.5, 4H, 2 x (-$C\underline{H}_2C(CH_3)_2O$-)), 3.14 (m, 6H, $C\underline{H}_2NHC(O)$, 2 x ($C\underline{H}_2NH(=NH)$)), 4.05 (m, 1H, C(α2)$\underline{H}$), 4.18 (t, *J*=7.0, 2H, C(9)$\underline{H}$), 4.29 (dd, *J*=5.5, 9.0, 1H, C(α1)$\underline{H}$), 4.40 (d, *J*=7.0, 2H, C(9)$HC\underline{H}_2$), 7.26 (t, *J*=7.5, 2H, C(2)$\underline{H}$, C(7)$\underline{H}$), 7.34 (t, *J*=7.5, 2H, C(3)$\underline{H}$, C(6)$\underline{H}$), 7.60 (d, br, *J*=7.5, C(1)$\underline{H}$, C(8)$\underline{H}$), 7.75 (d, *J*=7.5, 2H, C(4)$\underline{H}$, C(5)$\underline{H}$); **^{13}C NMR** (125.7MHz, CD_3OD) δ_C/ppm 12.5 (2 x C(e)$\underline{C}H_3$), 19.4, 19.6 (2 x C(b)$\underline{C}H_3$, 2 x C(f)$\underline{C}H_3$), 26.0, 27.5, 28.5, 29.9

($HO_2CCH_2\underline{C}H_2\underline{C}H_2\underline{C}H_2$, 2 x ($\underline{C}H_2\underline{C}H_2CH_2NH(=NH)$)), 36.1 ($HO_2C\underline{C}H_2$), 40.2, 41.4 ($\underline{C}H_2NHC(O)$, 2 x ($\underline{C}H_2NH(=NH)$)), 47.4 (2 x(-$\underline{C}H_2C(CH_3)_2O$-)), 49.5 ($\underline{C}$(9)), 54.3 ($\underline{C}(\alpha 1)$), 56.3 ($\underline{C}(\alpha 2)$), 68.3 (C(9)H$\underline{C}H_2$), 87.6 (2 x(-$CH_2\underline{C}(CH_3)_2O$-)), 118.4 (2 x $\underline{C}$(e)), 120.9 ($\underline{C}$(4), $\underline{C}$(5)), 126.1 ($\underline{C}$(1), $\underline{C}$(8)), 128.1 ($\underline{C}$(2), $\underline{C}$(7)), 128.7 ($\underline{C}$(3), $\underline{C}$(6)), 133.5 (2 x $\underline{C}$(b), 2 x $\underline{C}$(f)), 134.4 (2 x $\underline{C}$(a)), 139.4 ($\underline{C}$(4'), $\underline{C}$(5')), 142.5 ($\underline{C}$(1'), $\underline{C}$(8')), 145.2 (2 x $\underline{C}$(c)), 158.1 (2 x $\underline{C}$(d)), 158.5, 159.8, (NH$\underline{C}$(O)O, 2 x (NH($\underline{C}$=N))), 173.6, 174.4, 174.6 ($\underline{C}O_2H$, 2 x ($\underline{C}$(O)NH))); ν_{max} (nujol/cm^{-1}) 640(w), 663(w), 721(m), 759(m), 783(m), 806(m), 852(m), 968(w), 991(w), 1087(s), 1153(s, -SO_2-N), 1242(s), 1299(m), 1377(s, -SO_2-N), 1458(s), 1539(s, C(O)NH), 1543(s, C(O)NH), 1620(s, C=O), 1675(s, C=O), 1712(s, C=O), 2677(m, CO_2H), 2723(m, CO_2H), 2854(s, nujol) 2923(s, nujol), 3163(m, C(O)NH), 3317(s, C(O)NH), 3433(m, C(O)NH); **m/z** (FAB) 1193 (100%, $[M+Na]^+$), 1170 (60%, $[M+H]^+$); **HRMS** found 1192.5187, $[M+Na]^+$ ($C_{59}H_{79}O_{12}N_9S_2Na$) requires 1192.5160.

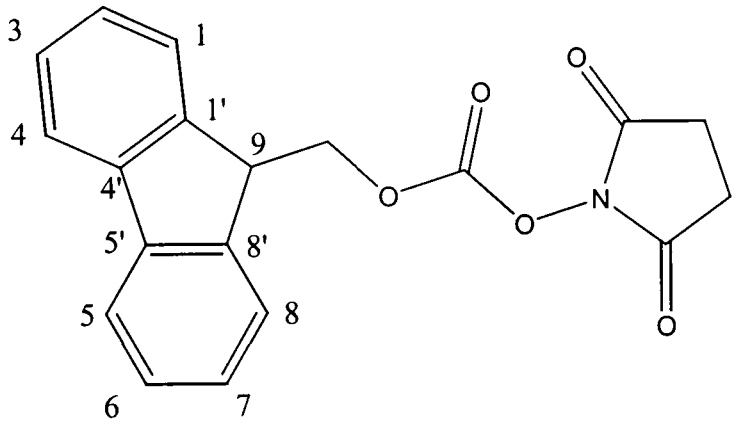
5. Synthesis of Polymers.

General: The polymer used was polyallylamine hydrochloride mw 70,000. The molecule weight of the monomer is 93, and number of mols. refers to the number of monomer equivalents in the polymer.

For spectral editing purposes we have used PFG BPP-LED NMR when analysing polymers. This is based on the fact that translational diffusion rates in solution for high and low molecular weights are significantly different (x100). Hence at high gradient strength (50%, 25Gcm^{-1}) signals corresponding to low molecular weight species will be suppressed, and thus the signals of the polymer observed more clearly. From experiments with varying gradient strengths we found that 15-20Gcm^{-1} was sufficient for suppression of peaks corresponding to species of molecular weight < 500amu.

5.0 Synthesis of Fmoc Protected Lysine (194).

Synthesis of 9-Fluorenylmethyl-Succinimidyl-Carbonate (Fmoc-OSu, 191)[146].

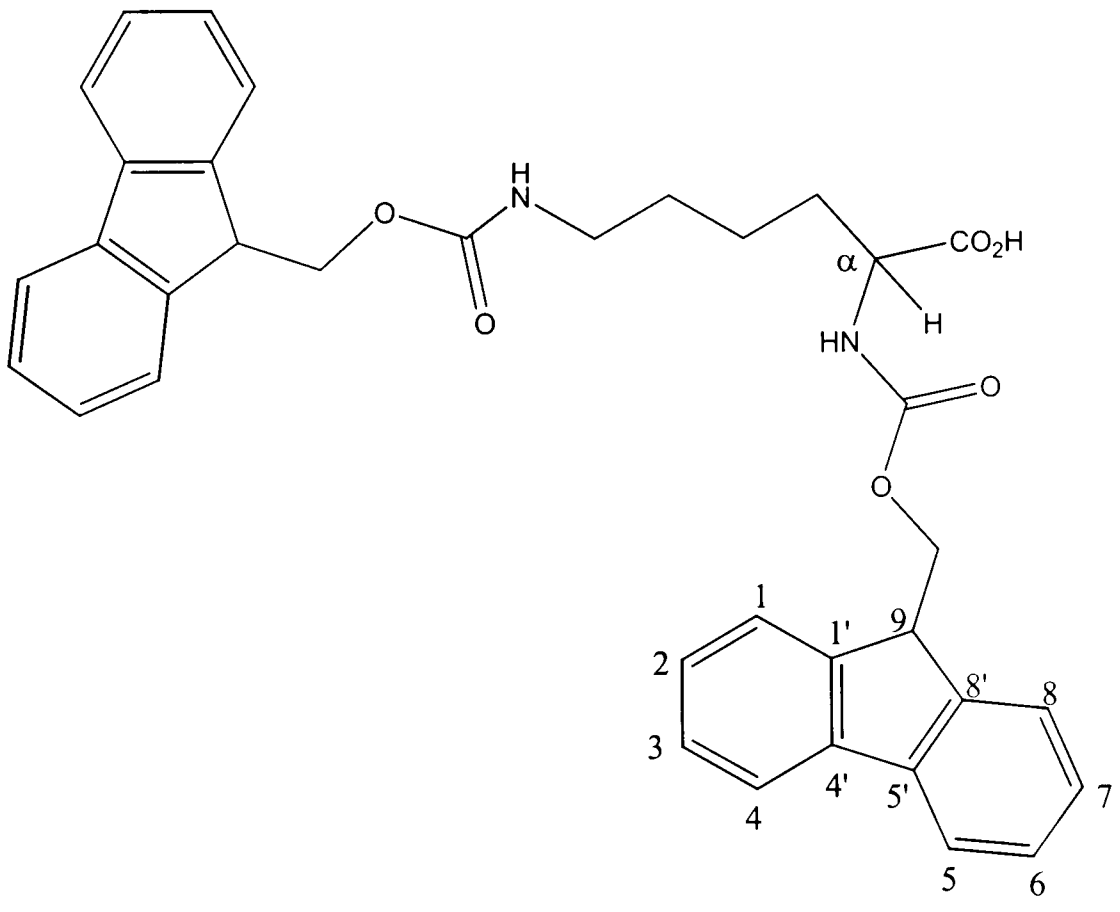


191

Triethylamine (2.65ml, 0.019mol) was added dropwise to a stirred solution of 9-fluorenylmethyl chloroformate (5g, 0.019mol) and *N*-hydroxysuccinimide (2.45g, 0.021mol) in dioxane (40ml) at 0°C. A white precipitate formed. The mixture was allowed to warm to r.t. and stirred for 1 hour. The white precipitate was removed by filtration, and the solvent removed *in vacuo* to yield a white solid. The solid was triturated with PE 30-40 (2 x 20ml) and dried under vacuum to yield **191** as a white solid (5.75g, 90%); **m.p.** 125-128°C (lit. 148-150°C[146]); ^{1}H **NMR** (500MHz,

$CDCl_3$) δ_H/ppm 2.81 (s, 4H, C(O)C$\underline{H}_2$C$\underline{H}_2$C(O)N), 4.33 (t, *J*=7.5, 1H, C(9)$\underline{H}$), 4.55 (d, *J*=7.5, 2H, C(9)HC$\underline{H}_2$), 7.33 (dd, *J*=7.5, 7.5, 2H, C(2)$\underline{H}$, C(7)$\underline{H}$), 7.42 (dd, *J*=7.5, 7.5 2H, C(3)$\underline{H}$, C(6)$\underline{H}$), 7.61 (d, *J*=7.5, 2H, C(1)$\underline{H}$, C(8)$\underline{H}$), 7.76 (d, *J*=7.5, 2H, C(4)$\underline{H}$, C(5)$\underline{H}$); **^{13}C NMR** (75.4MHz, $CDCl_3$) δ_C/ppm 25.7, 25.8 (C(O)$\underline{C}H_2\underline{C}H_2$C(O)N), 46.8 ($\underline{C}$(9)HCH$_2$), 73.2 (C(9)H$\underline{C}H_2$), 120.1, 120.5 ($\underline{C}$(4), $\underline{C}$(5)), 125.6 ($\underline{C}$(1), $\underline{C}$(8)), 127.5, 127.8 ($\underline{C}$(2), $\underline{C}$(7)), 128.6, 129.2 ($\underline{C}$(3), $\underline{C}$(6)), 141.7, 142.9 ($\underline{C}$(1'), $\underline{C}$(4'), $\underline{C}$(5'), $\underline{C}$(8')), 151.9 (O$\underline{C}$(O)O), 169.1, 172.3 (2 x N$\underline{C}$(O)); **ν_{max}** (nujol/cm^{-1}) 709(w), 731(m), 756(m), 935(w), 950(m), 989(m), 1074(w), 1157(s), 1201(m), 1230(m), 1263(m), 1377(m), 1461(s, nujol), 1714.6(s), 1737.7(s, C=O), 1784.0(s, C=O), 1814.9(m), 2339.5(m), 2678.9(m), 2852.5(s, CH_2, nujol), 2922.0 (s, CH_2, nujol); **m/z** (FAB) 470 (100%), 379 (36%), 337 (34%, $[M]^+$); **HRMS** found 337.0950, $[M]^+$ ($C_{19}H_{15}O_5N$) requires 337.0938.

Synthesis of 2,6-*bis*-(*9*-H-Fluoren-9-ylmethoxycarbonylamino)-Hexanoic Acid (FmocLys(Fmoc)OH, 194)[164].

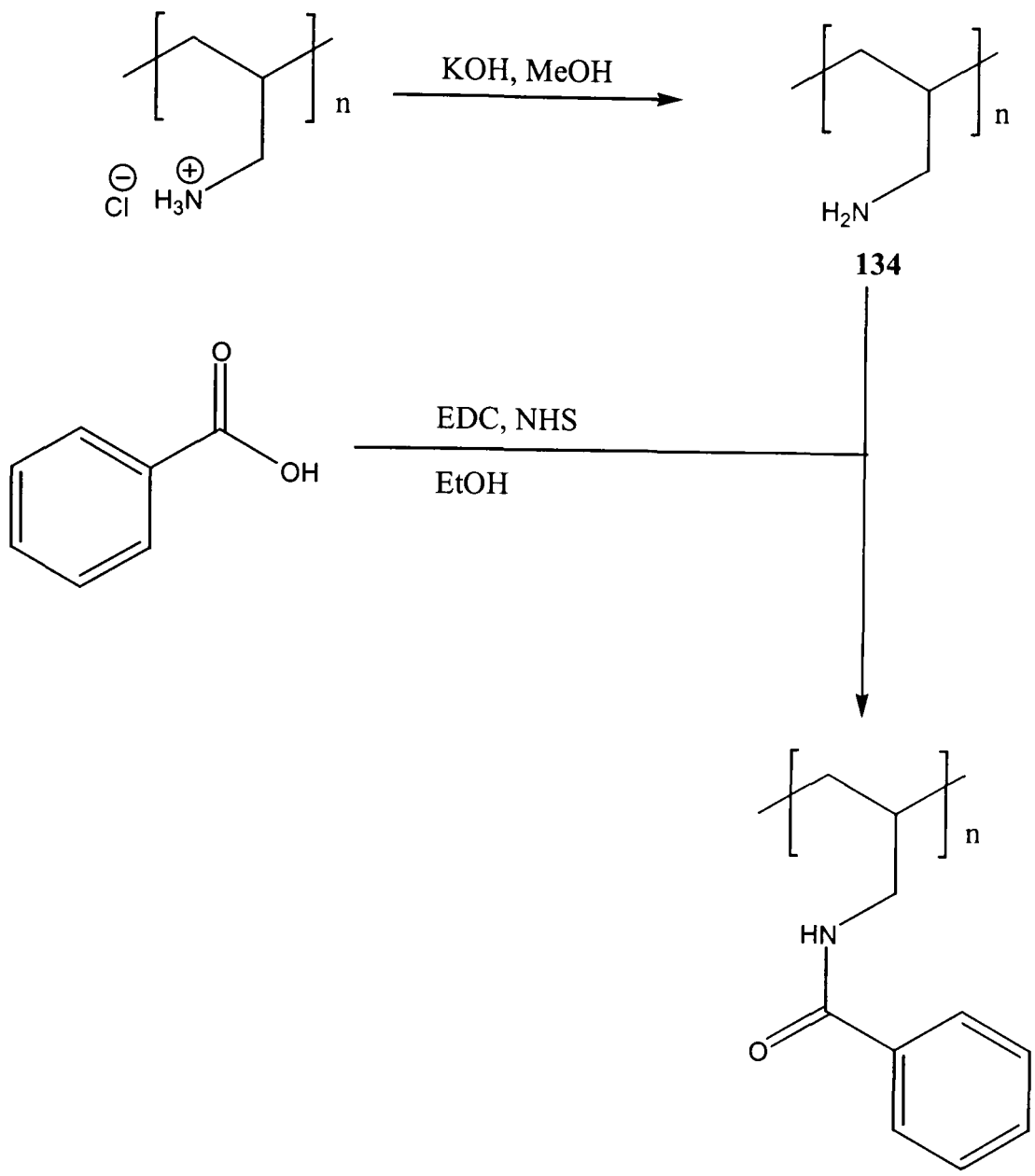


194

Lysine.HCl (1.35g, 7.4mmol) was dissolved in aq. Na_2CO_3 solution (9%, 2.36g, 22mmol in 26ml). The solution was cooled to 0°C, and to it added Fmoc-OSu **191** (5g,

15mmol) in dioxane (30ml). The solution was allowed to warm to r.t. and stirred for 1 hour. After this time a white precipitate had formed. The mixture was diluted with H_2O (400ml), and washed with Et_2O (1 x 100ml), and EtOAc (1 x 100ml). The pH of the aqueous layer was adjusted to 1 using conc. HCl. A white precipitate formed which was extracted into $CHCl_3$ (3 x 150ml). The organic layers were combined, dried over Na_2SO_4, and the solvent removed *in vacuo* to yield a cream oil, which upon cooling solidified. The solid was suspended in water, stirred for 10 hours, and filtered to yield **194** as a white solid (4.03g, 71%); **m.p.** 118°C; $\mathbf{R_f}$ 0.47 (SiO_2; EtOAc:MeOH; 8:2); **^{1}H NMR** ($CDCl_3$, 500MHz) δ_H/ppm 1.23-1.72 (m, 6H, C$\underline{H}_2$C$\underline{H}_2$C$\underline{H}_2$CH(CO_2H)NH), 3.55 (m, br, 2H, C$\underline{H}_2$NHC(O)), 4.15 (s, br, 2H, C(9)$\underline{H}$), 4.33 (m, 4H, C(9)HC$\underline{H}_2$), 4.44 (m, 1H, C(α)$\underline{H}$), 4.96 (s, br, 2H, 2 x N$\underline{H}$), 7.24 (t, *J*=7.5, 4H, C(3)$\underline{H}$, C(6)$\underline{H}$), 7.33 (t, *J*=7.5, 4H, C(2)$\underline{H}$, C(7)$\underline{H}$), 7.54 (d, *J*=7.5, 4H, C(1)$\underline{H}$, C(8)$\underline{H}$), 7.70 (d, *J*=7.5, 4H, C(4)$\underline{H}$, C(5)$\underline{H}$); **^{13}C NMR** ($CDCl_3$, 125.7MHz) δ_C/ppm 22.2 (C$\underline{H}_2$CH$_2$C(α)), 29.2 (C$\underline{H}_2$CH$_2$NHC(O)O), 31.7 (C$\underline{H}_2$C(α)), 40.4 (C$\underline{H}_2$NHC(O)O), 47.1 (2 x C(9)), 53.6 (C(α)), 66.8 (2 x C(9)HC$\underline{H}_2$), 119.9 (C(4), C(5)), 124.8, 124.9 (C(1), C(8)), 126.5 (C(2), C(7)), 127.6 (C(3), C(6)), 141.2, 143.6 (C(1'), C(4'), C(5'), C(8')), 156.3, 156.7 (2 x OC(O)NH), 175.7 ($\underline{C}O_2$H); ν_{max} (nujol/cm^{-1}) 732(m), 757(m), 1038(m), 1083(m), 1103(m), 1148(m), 1234(m), 1264(s), 1339(m), 1374(m), 1454(s, nujol), 1535(s, OC(O)NH), 1655(m), 1690(s, C=O), 1725(s, C=O), 2675(m, CO_2H), 2725(m, CO_2H), 2845(s, nujol), 2925(s, nujol), 3066(m, OC(O)NH), 3317(s, OC(O)NH); **m/z** (FAB) 613 (100%, $[M+Na]^+$), 591 (42%, $[M+H]^+$); **HRMS** found 591.2510, $[M+H]^+$ ($C_{36}H_{35}N_2O_6$) requires 591.2495.

5.1 Synthesis of Polymer containing Benzoic Acid (201).



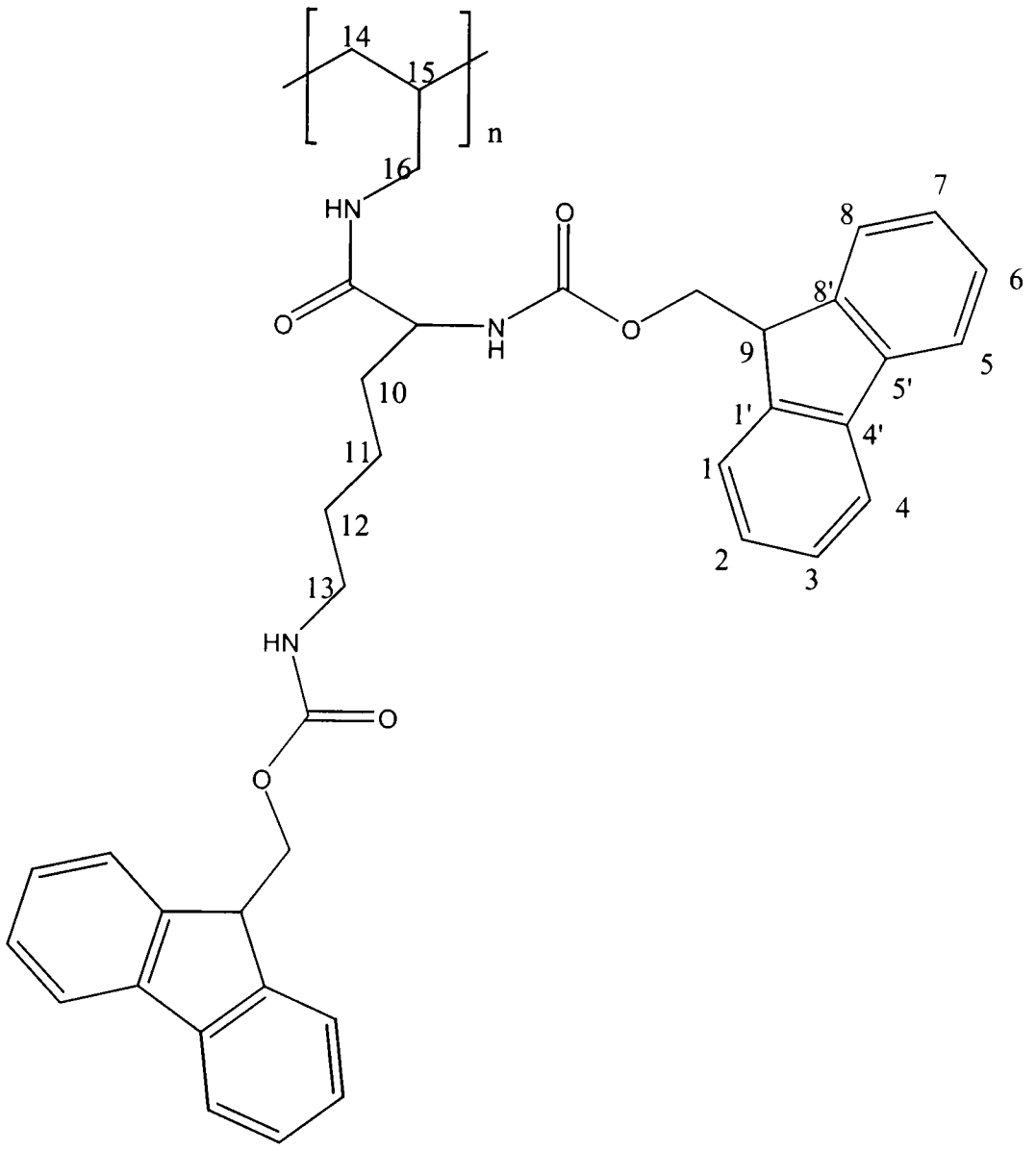
201

Polyallylamine hydrochloride (200mg, 2.14mmol) was stirred with potassium hydroxide (100mg, 2.2mmol) in MeOH (30ml) for 15 hours. The solvent was removed *in vacuo*, and the residue taken up in EtOH (30ml). A white solid (KCl) did not dissolve. The solution was filtered, and concentrated to a volume of 10ml *in vacuo*, to yield a solution of free polyallylamine **134** in EtOH.

Benzoic Acid (643mg, 5.3mmol), *N*-hydroxysuccinimide (848mg, 7.4mmol) and EDC (1.2g, 6.3mmol) were dissolved in EtOH (60ml) and stirred at r.t. for 4 hours. The polyallylamine solution was added to the reaction. A white precipitate formed instantly. The solution was filtered, and the precipitate collected to yield the required polymer **201** as a white gum (956mg); **^{13}C NMR** (75.4MHz, solid state) 14.5, 18.6, 26.7, 33.5 ($\underline{C}H_2$, $\underline{C}H$ from polyallylamine), 57.8, 118.7-129.1 (Ar$\underline{C}$), 177.6 ($\underline{C}$(O)NH).

5.2 Synthesis of Polymer 204 Containing Lysine.

Reaction of Polyallylamine with FmocLys(Fmoc)-OH to Afford 203.



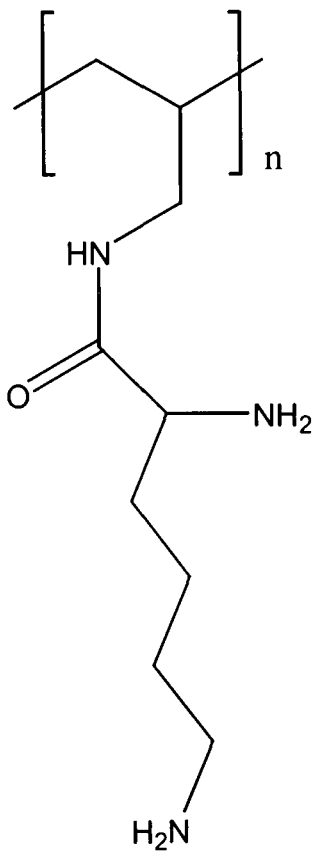
203

Polyallylamine hydrochloride (100mg, 1.07mmol) was stirred with potassium hydroxide (50mg, 1.1mmol) in MeOH (20ml) for 15 hours. The solvent was removed *in vacuo*, and the residue dissolved in EtOH (20ml). The solution was filtered to remove any undissolved KCl which had formed, and concentrated to 10ml *in vacuo*, to yield a solution of free polyallylamine in EtOH.

FmocLys(Fmoc)-OH **194** (874mg, 1.6mmol), DCC (396mg, 1.9mmol) and *N*-hydroxysuccinimide (258mg, 2.2mmol) were dissolved in $CHCl_3$ (50ml) and stirred at r.t. for 3 hours. The solution of polyallylamine was added to this solution, and the reaction stirred for a further 3 hours, after which time a white precipitate had formed.

The solid was collected by filtration, and dried under vacuum to yield the desired polymer **203** as a white insoluble solid (1.023g); **^{13}C NMR** (75.4MHz, solid state) δ_C/ppm 24.7 (C(11)), 27.7 (C(14)), 33.3 (C(10), C(12)), 36.2 (C(16)), 40.9 (C(15)), 42.1 (C(13)), 46.6 (C(9)), 53.8, 55.8 (CHC(O)NH), 61.9, 63.1, 67.7 (C(9)HCH$_2$), 120.6 (C(1), C(8)), 126.8, (C(2), C(3), C(4), C(5), C(6), C(7)), 141.9 (C(1'), C(4'), C(5'), C(8')), 157.4 (NHC(O)O), 173.7 (RC(O)NH); ν_{max} (nujol/cm^{-1}) 736(m), 756(m), 871(w), 937(w), 1033(m), 1087(m), 1103(w), 1141(m), 1242(s), 1377(s), 1461(s, nujol), 1539(s, NHC(O)), 1651(s, C=O), 1693(s, C=O), 2854(s, nujol), 2923(s, nujol), 3294(s, br).

Removal of Base Labile (Fmoc) Protecting Groups to Afford Polymer 204.

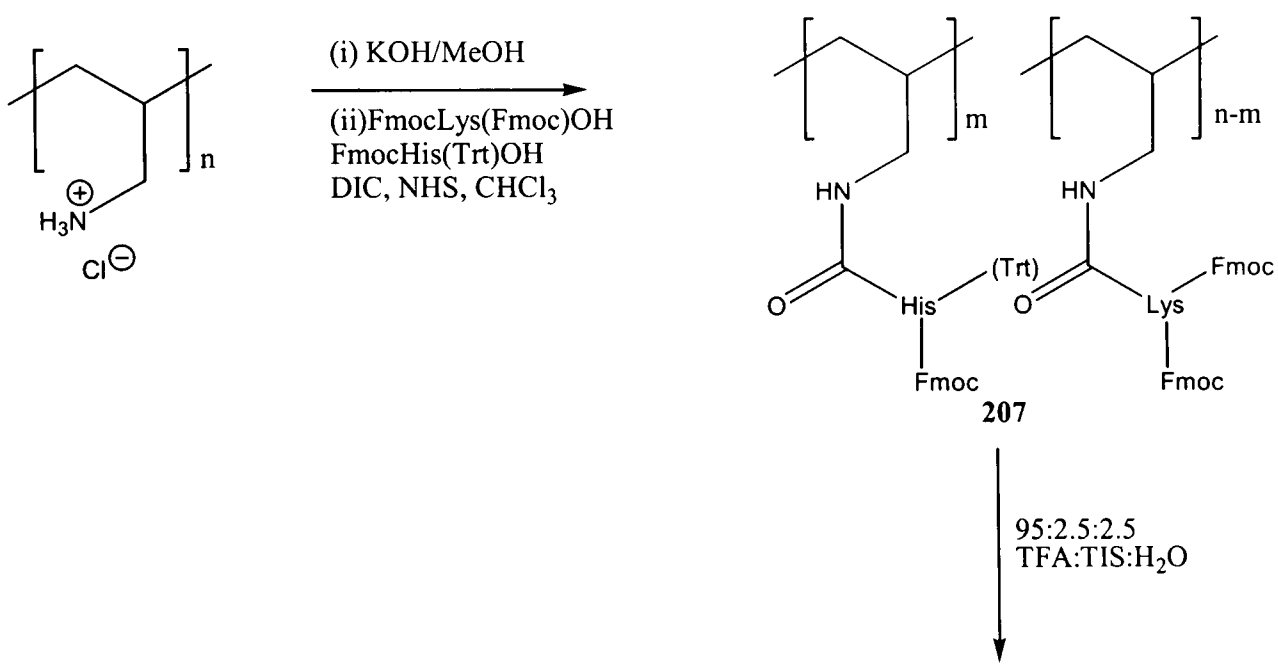


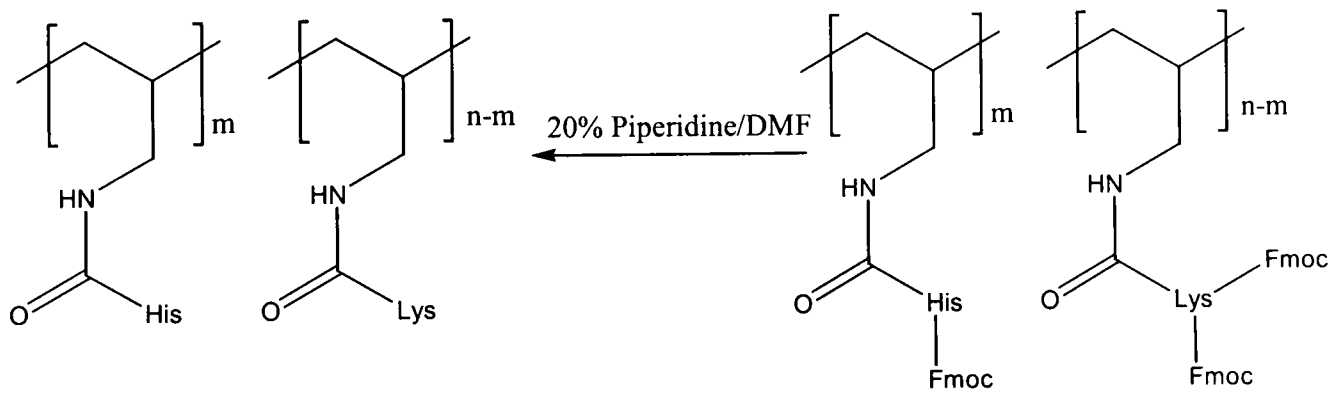
204

Polymer **203** (583mg) was suspended in a solution of 20% piperidine/DMF (25ml) and stirred for 2 hours. The mixture was filtered, and the solid dried under vacuum, to afford the desired polymer **204** as a white solid (85mg); **^{1}H NMR** (500MHz, D_2O) δ_H/ppm 0.82, 1.31 (CH$_2$CH(CH$_2$)CH$_2$NHC(O), CH$_2$(CH$_2$)$_2$NH$_2$), 1.54 (CH$_2$CH$_2$NH$_2$), 1.80 (CH$_2$CH(C(O)NH)NH$_2$), 2.89, 3.05 (CH$_2$CH(CH$_2$)CH$_2$NHC(O), CH$_2$NH$_2$, CH$_2$NHC(O)), 3.95 (CHC(O)NH); **^{13}C NMR** (125.7MHz, D_2O) δ_C/ppm 19.5 (CH$_2$CH(CH$_2$)CH$_2$NHC(O)), 23.3 (CH$_2$(CH$_2$)$_2$NH$_2$), 28.1 (CH$_2$CH(CH$_2$)CH$_2$NHC(O)), 32.4, 35.6 (CH$_2$CH(C(O)NH)NH$_2$, CH$_2$CH$_2$NH$_2$), 40.8 (CH$_2$NH$_2$), 53.5, 54.9

(CHC(O)NH, CH_2NHC(O)), 171.2 (CHC(O)NH); ν_{max} (solid state/cm^{-1}) 736(w), 818(m), 1088(m), 1243(w), 1310(m), 1369(m), 1452(s), 1562.41(s, C(O)NH), 1642(s, C=O), 2854(m), 2932(m, C(O)NH), 3274(m, C(O)NH).

5.3 Synthesis of Polymer 217 Containing Lysine and Histidine.





Reaction of Polyallylamine with FmocLys(Fmoc)-OH, and FmocHis(Trt)-OH to Afford 207.

Polyallylamine hydrochloride (200mg, 2mmol) was stirred with potassium hydroxide (96mg, 2mmol) in MeOH (15ml) for 13 hours. The solvent was removed *in vacuo*, and

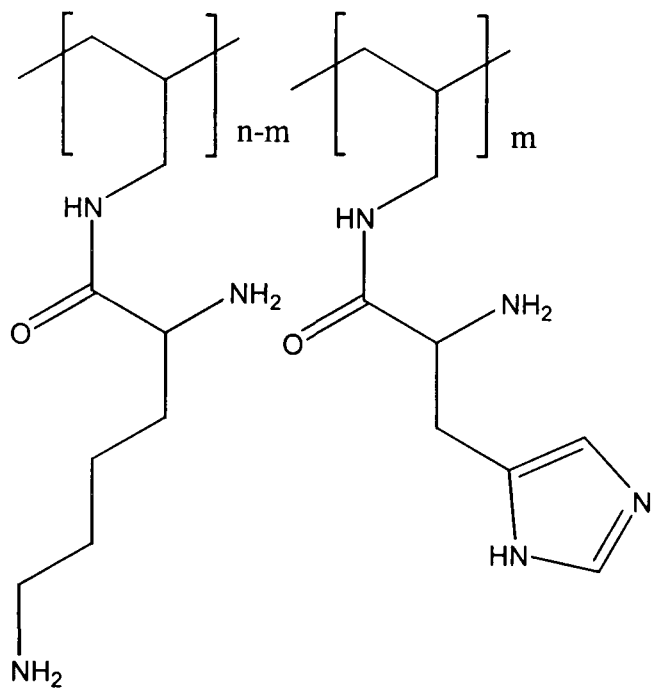
the residue dissolved in EtOH (20ml). The solution was filtered, and concentrated to a volume of 5ml under reduced pressure, to yield a solution of free polyallylamine in EtOH.

FmocHis(Trt)-OH (533mg, 0.86mmol), FmocLys(Fmoc)-OH (761mg, 1.3mmol), DIC (376µl, 2.4mmol) and *N*-hydroxysuccinimide (322mg, 2.8mmol) were dissolved in $CHCl_3$ (100ml) and stirred for 3 hours, until no further reaction was observed by TLC (SiO_2; EtOAc:MeOH; 9:1). The solution of free polyallylamine in EtOH was then added to the reaction, and the resulting solution left to stir for 24 hours. After this time a white precipitate had formed which was removed by filtration. The solution was concentrated to 30ml, and diluted with MeOH (100ml). A cream precipitate formed, which was collected by filtration, and washed with DCM: MeOH (2:1, 2 x 40ml), and dried under vacuum to yield the desired polymer **207** (975mg); **^{1}H NMR** (500MHz, $CDCl_3$ - See Appendix 1 for spectrum) δ_H/ppm 1.0-1.9 (br, C$\underline{H}_2$ of polyallylamine and amino acids), 2.80, 3.20 (br, C$\underline{H}$ of polyallylamine, C$\underline{H}_2$N of amino acids), 3.80-4.40 (br, α$\underline{H}$ of amino acids), 6.80-7.90 (br, $\underline{H}$ of Fmoc and Trityl groups); ν_{max} (solid state/cm^{-1}) 670(m), 720(m), 749(w), 892(w), 1047(w), 1088(m), 1158(w), 1186(w), 1243(m), 1319(m), 1403(m), 1437(m), 1571(s, C(O)NH), 1625(s, C=O), 1667(s, C=O), 2851(m), 2928(m), 3324(m).

Deprotection of Acid Labile (Trt) Protecting Groups to Afford 220.

Polymer **207** (975mg) was suspended in a solution of TFA:triisopropylsilane:H_2O (95:2.5:2.5, 30ml) and sonicated for 4 hours. The mixture was filtered, and the solid washed with H_2O (2 x 30ml), Et_2O (2 x 30ml), H_2O (1 x 30ml) and dried under vacuum to yield the desired polymer **220** (531.5mg); **^{1}H NMR** (500MHz, $CDCl_3$ - See Appendix 1 for spectrum) δ_H/ppm 1.05-1.61 (br, C$\underline{H}_2$ of polyallylamine/lysine), 2.7-3.5 (br, C$\underline{H}_2$ of histidine, C$\underline{H}$ of polyallylamine, C$\underline{H}_2$NH- of lysine), 3.7-4.19 (br, α$\underline{H}$ of peptides), 6.50 (br, C$\underline{H}$=CR_2), 7.09-7.78 (br, Fmoc$\underline{H}$).

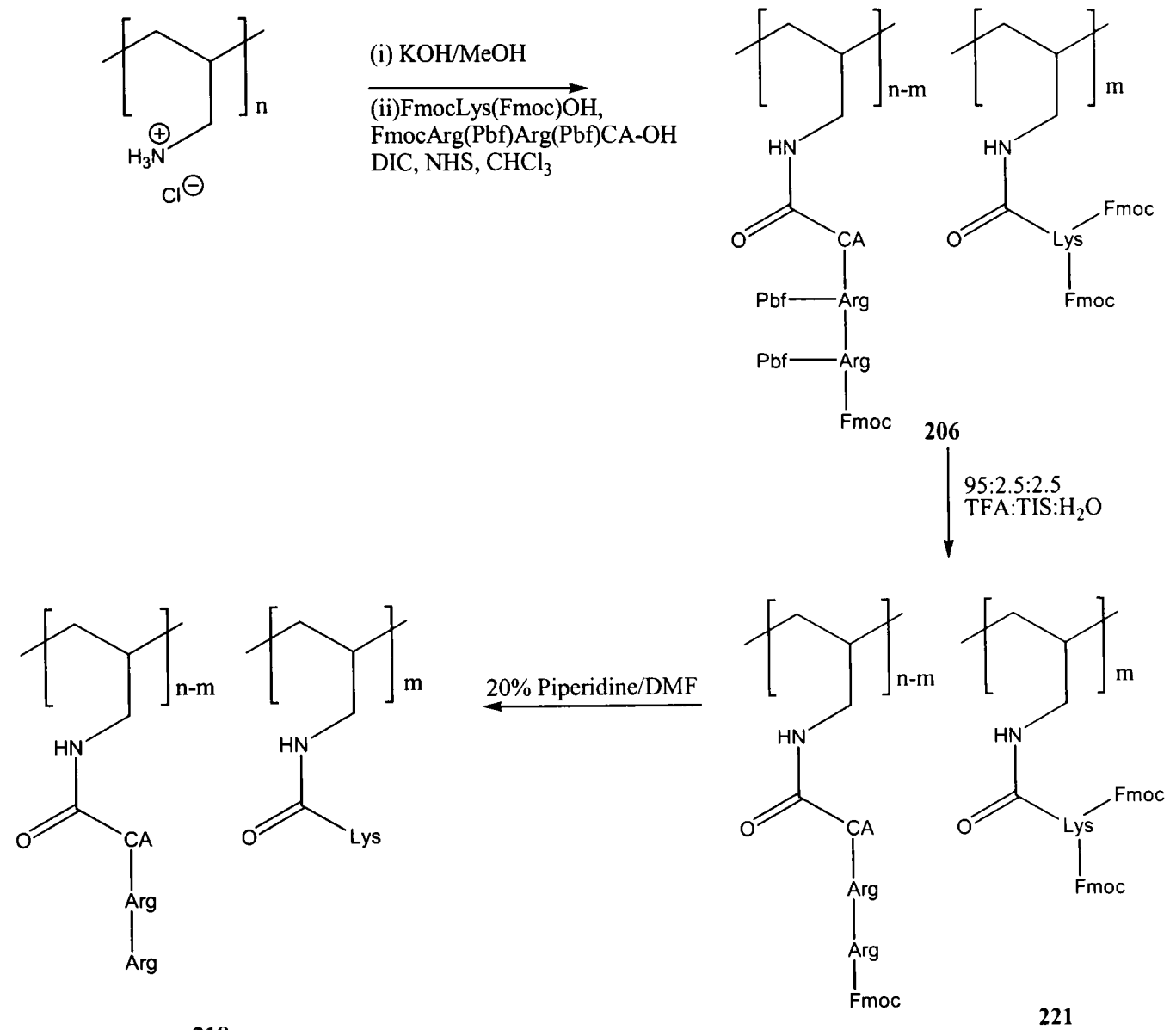
Removal of Base Labile (Fmoc) Protecting Groups to Afford Polymer 217.



217

Polymer **220** (531mg) was suspended in 20% piperidine/DMF (25ml) and sonicated for 4 hours. After this time the mixture was filtered. The solid was washed with H_2O (1 x 40ml), DMF (2 x 30ml), H_2O (1 x 20ml), DCM (1 x 30ml), and dried *in vacuo* to yield the desired polymer **217** (93mg). **^{1}H NMR** (500MHz, D_2O – See Appendix 1 for spectrum) δ_H/ppm 1.14, 1.40, 1.65, 1.85 (br, C$\underline{H}_2$ of lysine/polyallyamine), 2.41, 2.93, 3.27 (C$\underline{H}$ of polyallyalmine, C$\underline{H}_2$NH$_2$ of lysine, C$\underline{H}_2$ of histidine), 4.38, 4.76 ($\alpha\underline{H}$ of peptides), 7.68, 8.67 (C$\underline{H}$=N, C$\underline{H}$=CR$_2$, histidine ring).

5.4 Synthesis of Polymer 218 Containing Lysine and Arginine-Arginine.



Reaction of Polyallylamine with FmocArg(Pbf)Arg(Pbf)CA-OH and FmocLys(Fmoc)-OH to Afford 206.

Polyallylamine hydrochloride (150mg, 1.4mmol) was stirred with potassium hydroxide (100mg, 2mmol) in MeOH (50ml) for 14 hours. The solvent was removed *in vacuo*, and the residue dissolved in EtOH (50ml). The solution was filtered to remove any KCl formed, and the solution concentrated to 10ml, yielding a solution of free polyallylamine in EtOH.

FmocArg(Pbf)Arg(Pbf)CA-OH (622mg, 0.5mmol), DIC (303µl, 1.9mmol) and *N*-hydroxysuccinimide (260mg, 2.2mmol) were dissolved in $CHCl_3$ (50ml) and stirred at

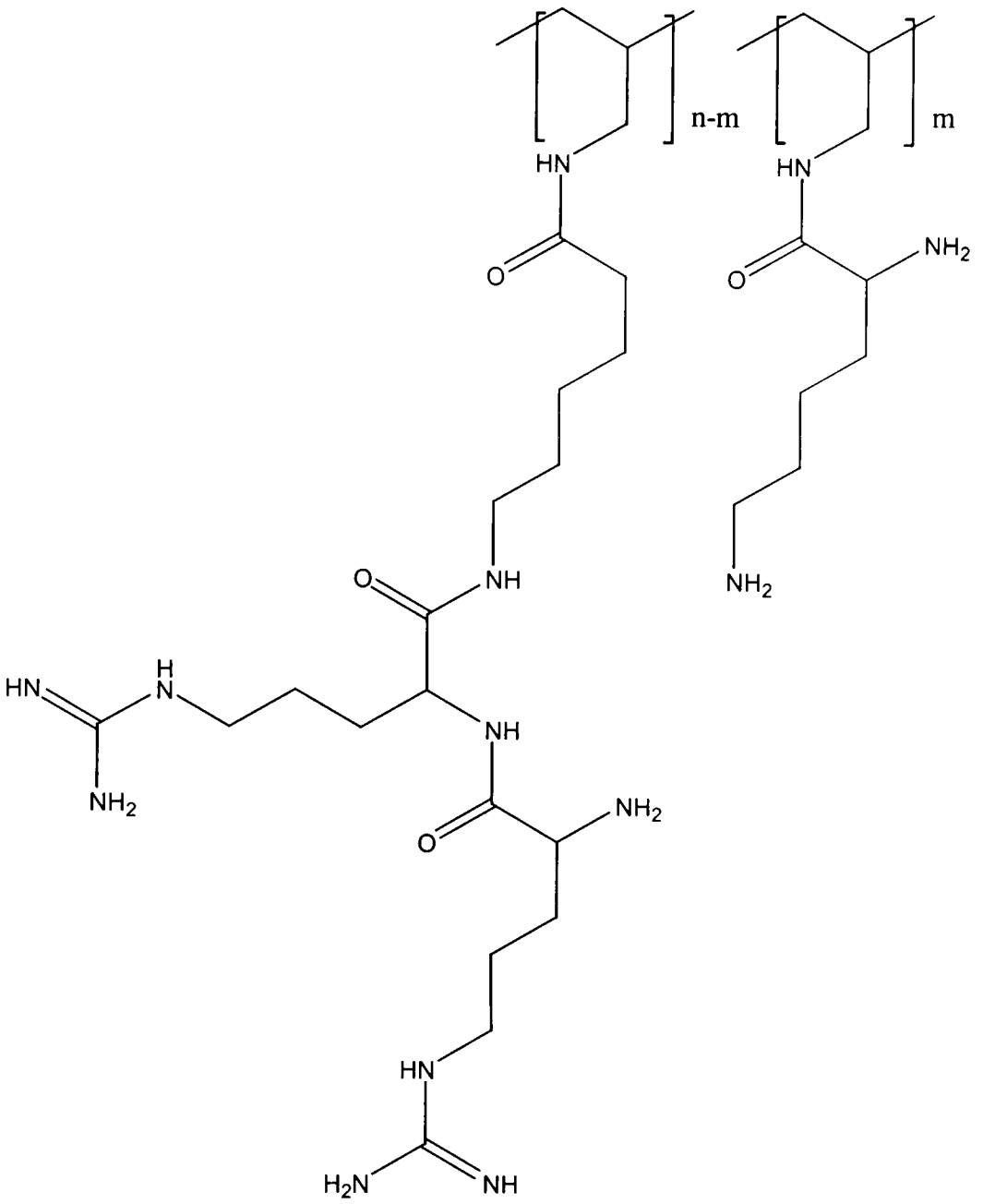
r.t. for 3 hours, until TLC (SiO_2; EtOAc:MeOH; 9:1) showed no further reaction. The ethanolic solution of polyallylamine was then added, and the mixture stirred at r.t.

At the same time, FmocLys(Fmoc)-OH (623mg, 1mmol), DIC (198µl, 1.3mmol) and *N*-hydroxysuccinimide (170mg, 1.5mmol) were dissolved in $CHCl_3$ (50ml) and stirred for 2 hours. This solution was then added to the reaction containing polyallylamine. The reaction was stirred for 48 hours. After this time a white precipitate had appeared which was removed by filtration. The solution was concentrated *in vacuo* to 20ml, and diluted with MeOH (100ml). A white gum formed, which was collected by filtration, and washed with $CHCl_3$ (30ml), and MeOH (50ml). The gum was dried under to vacuum to yield the desired polymer **206** (564mg); **^{1}H NMR** (500MHz, CD_3OD - see appendix 1 for spectrum) δ_H/ppm 1.16-1.41 (br, C$\underline{H}_2$ from polyallylamine/ amino acids), 2.04 (br, ArC$\underline{H}_3$ of Pbf group), 2.57, 3.23, 3.87-4.39 (br, α$\underline{H}$ of peptides), 7.32-7.77 (br, Fmoc$\underline{H}$).

Cleavage of Acid Labile (Pbf) Protecting Groups to Afford 221.

Polymer **206** (560mg) was suspended in a solution of TFA:triisopropylsilane:H_2O (95:2.5:2.5, 40ml) and sonicated at r.t. for 4 hours. The mixture was filtered, and washed with H_2O (1 x 50ml), TFA (1 x 30ml), Et_2O (1 x 30ml), $CHCl_3$ (1 x 40ml) and H_2O (1 x 50ml), and dried under vacuum to yield the desired polymer **221** as a yellow gum (249mg); **^{1}H NMR** (500MHz, CD_3OD- see appendix 1 for spectrum) δ_H/ppm 1.19-2.20 (br, C$\underline{H}_2$ of polyallylamine/amino acids), 2.46, 3.17, 4.13-4.35 (br, α$\underline{H}$ of peptides), 7.27-7.71 (br, Fmoc$\underline{H}$); ν_{max} (solid state/cm^{-1}) 749(s), 758(m), 1137(m), 1202(m), 1246(m), 1449(w), 1508(m), 1524(m), 1541(s), 1653(s, C=O), 1685(s, C=O), 2941(m), 3237(m, C(O)NH), 3343(m, C(O)NH).

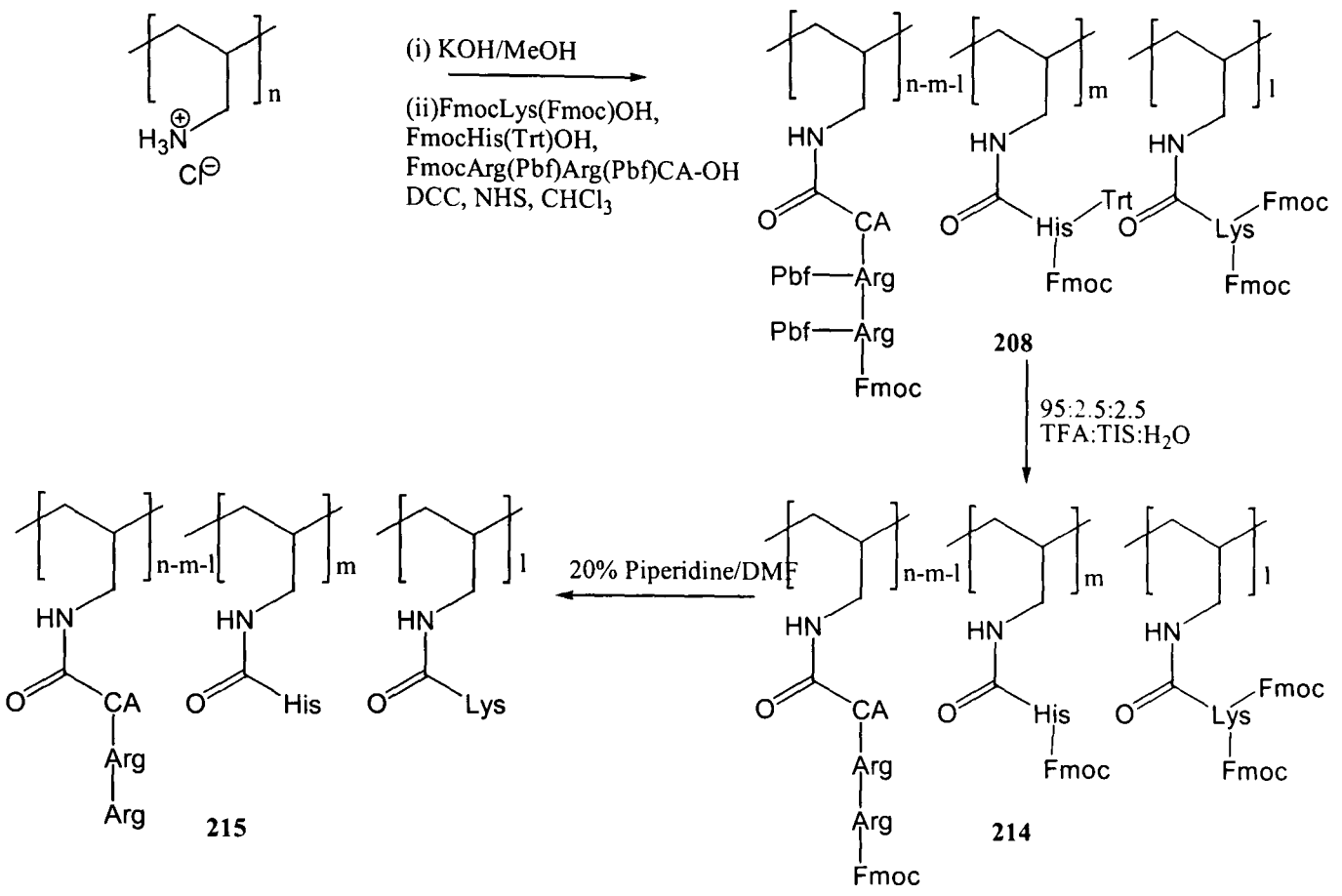
Cleavage of Base Labile (Fmoc) Protecting Groups to Afford Polymer 218



218

Polymer **221** (240mg) was suspended in a solution of 20% piperidine/DMF (40ml) and sonicated for 5 hours. The mixture was filtered and the polymer product was removed from the filter paper by sonicating in H_2O (30ml) for 5 hours. The solvent was removed *in vacuo*, to yield the fully deprotected polymer **218** as a yellow-brown solid (79.8mg); **^{1}H NMR** (500MHz, D_2O- see appendix 1 for spectrum) δ_H/ppm 1.15-1.96 (br, C$\underline{H}_2$ of polyallylamine/amino acids) 2.14, 2.26, 2.47, 2.86-3.60 (br, C$\underline{H}$ of polyallylamine, C$\underline{H}_2$NH of amino acids) 3.93- 4.16 (br, $\alpha\underline{H}$ of amino acids).

5.5 Synthesis of Polymer 215 Containing Lysine, Histidine and Arginine-Arginine.



Reaction of Polyallylamine with FmocArg(Pbf)Arg(Pbf)CA-OH, FmocLys(Fmoc)-OH and FmocHis(Trt)-OH to Afford 208.

Polyallylamine hydrochloride (121mg, 1.3mmol) was stirred in MeOH (20ml) with potassium hydroxide (50mg, 1.5mmol) for 16 hours. The solvent was removed *in vacuo*, and the residue dissolved in EtOH (20ml). The solution was filtered to remove any KCl formed during the reaction, and concentrated to 10ml to yield a solution of free polyallylamine.

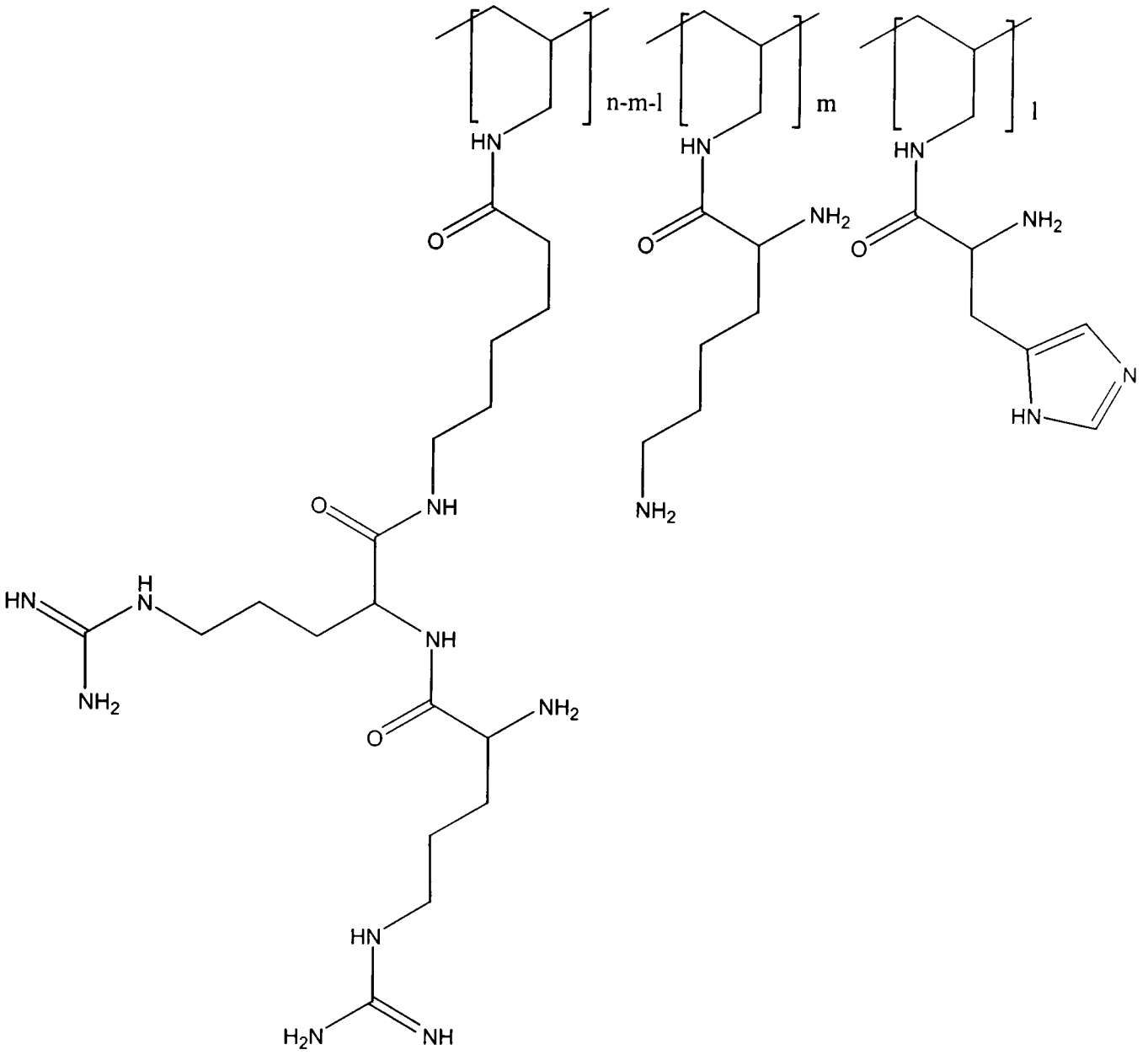
FmocArg(Pbf)Arg(Pbf)CA-OH (450mg, 0.38mmol), FmocHis(Trt)-OH (235mg, 0.38mmol), DCC (188mg, 0.91mmol) and *N*-hydroxysuccinimide (122mg, 1.06mmol) were dissolved in $CHCl_3$ (50ml) and stirred for 3 hours until TLC (SiO_2; EtOAc:MeOH; 9:1) showed no further reaction. The solution of polyallylamine was added, and the resultant solution stirred for 2 hours.

At the same time, FmocLys(Fmoc)-OH (383mg, 0.65mmol), DCC (161mg, 0.78mmol) and *N*-hydroxysuccinimide (105mg, 0.91mmol) were dissolved in $CHCl_3$ (40ml) and stirred. After 2 hours, this solution was added to the reaction containing polyallylamine. The resulting solution was stirred for 24 hours. After this time the solution was diluted with MeOH (100ml). A precipitate formed which was collected by filtration, and dried under vacuum to afford the desired polymer **208** as a yellow gel (333mg); **^{1}H NMR** (500MHz, $CDCl_3$ – see appendix 1 for spectrum) δ_H/ppm 0.91-1.80 (br, C$\underline{H}_2$ of polyallylamine and amino acids), 2.04 (br, ArC$\underline{H}_3$ of Pbf), 2.50, 3.09 (br, C$\underline{H}$ of polyallylamine, C$\underline{H}_2$N of amino acids), 3.81-4.25 (br, $\alpha\underline{H}$ of amino acids), 7.21-7.77 (br, $\underline{H}$ of Fmoc and Trityl groups); ν_{max} (solid state/cm^{-1}) 732(m), 810(w), 1023(m), 1090(s), 1149(w), 1257(m), 1301(m), 1386(m), 1555(m), 1656(s, C=O), 2935(m), 3246(m, C(O)NH), 3312(m, C(O)NH).

Cleavage of Acid Labile (Pbf, Trt) Protecting Groups to Afford 214.

Polymer **208** (300mg) was suspended in a solution of TFA:triisopropylsilane:H_2O (95:2.5:2.5, 15ml) and sonicated at r.t. for 3 hours. The mixture was filtered, and washed with H_2O (2 x 40ml), TFA (1 x 50ml), and H_2O (1 x 50ml), and dried under vacuum to yield the desired polymer **214** as a yellow solid (227mg); **^{1}H NMR** (500MHz, CD_3OD – see appendix 1 for spectrum) δ_H/ppm 1.21-1.90 (br, C$\underline{H}_2$ of polyallylamine/ amino acids), 2.47, 2.93, 3.20 (C$\underline{H}$ of polyallylamine, C$\underline{H}_2$NH of amino acids), 3.8-4.58 (br, $\alpha\underline{H}$ of amino acids), 7.26-7.71 (br, Fmoc$\underline{H}$); ν_{max} (solid state/cm^{-1}) 652(w), 725(s), 799(m), 823(w), 907(m), 1135(s), 1199(s), 1448(w), 1549(m), 1670(s, C=O), 2917(m), 3089(m, br, C(O)NH).

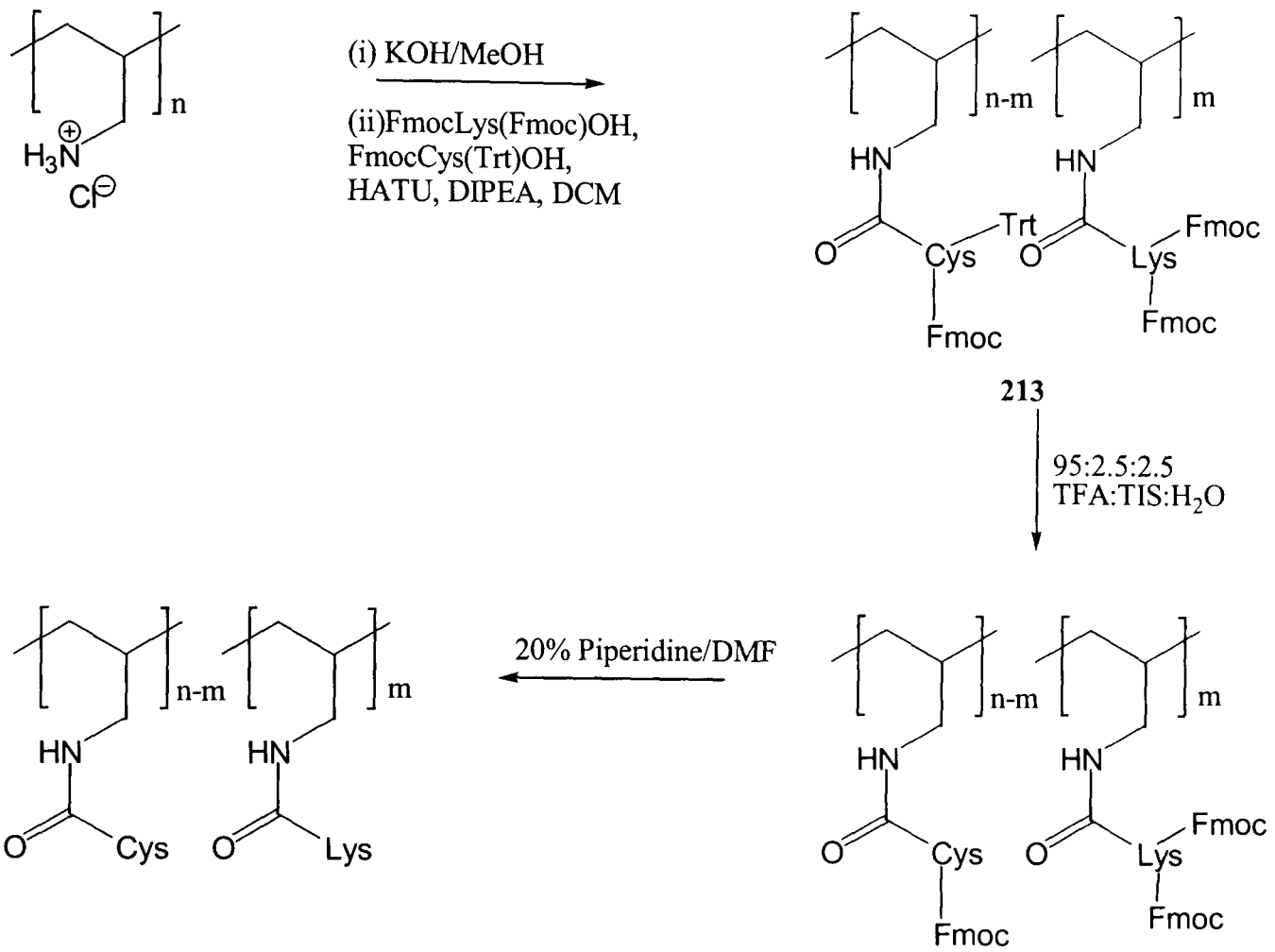
Cleavage of Base Labile (Fmoc) Protecting Groups to Afford Polymer 215.



215

Polymer **214** (220mg) was suspended in a 20% piperidine/DMF solution (40ml) and sonicated at r.t. for 3 hours. The mixture was filtered and washed with DMF (3 x 40ml), H_2O (2 x 40ml) and DCM (2 x 40ml). The solid was dried under vacuum to yield the fully deprotected polymer **215** as an orange solid (84mg); **^{1}H NMR** (500MHz, CD_3OD – see appendix 1 for spectrum) δ_H/ppm 0.9-2.1 (br, C<u>H</u>$_2$ of polyallylamine and amino acids), 2.2-3.5 (br, C<u>H</u> of polyallylamine, C<u>H</u>$_2$ of histidine, C<u>H</u>$_2$NH of lysine/arginine), 3.8-4.3 (br, α<u>H</u> of amino acids), 7.32 (br, C<u>H</u>=N, C<u>H</u>=CR_2 of histidine); ν_{max} (solid state/cm^{-1}) 673(s), 727(m), 755(m), 1070(s), 1995(s), 1151(m), 1214(m), 1232(m), 1384(m), 1447(m), 1561(s, br), 1644(s, C=O), 2929(m), 2942(m), 3189(m, C(O)NH), 3299(m, C(O)NH), 3367(m, C(O)NH).

5.6 Synthesis of Polymer 216 Containing Lysine and Cysteine.



216 **213** **222**

Reaction of Polyallylamine with FmocCys(Trt)-OH and FmocLys(Fmoc)-OH to Afford 213.

Polyallylamine hydrochloride (200mg, 2mmol) was stirred with potassium hydroxide (100mg, 2mmol) in MeOH (50ml) for 14 hours. The solvent was removed *in vacuo*, and the residue dissolved in EtOH (50ml). The solution was filtered to remove any KCl formed, and the solution concentrated to 10ml, yielding a solution of free polyallylamine in EtOH.

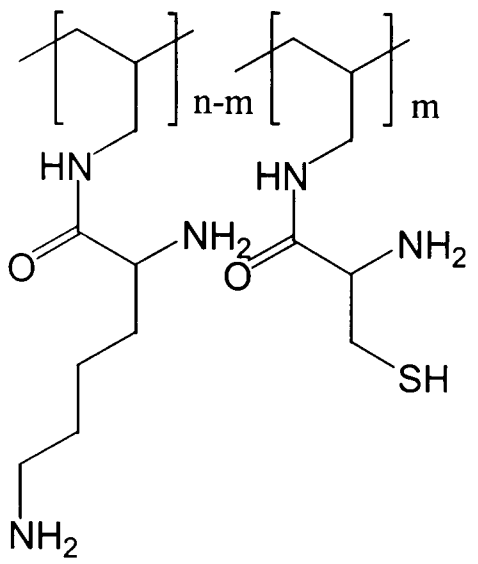
FmocCys(Trt)-OH (470mg, 0.8mmol), HATU (304mg, 0.8mmol) and DIPEA (279ul, 1.6mmol) were dissolved in DCM (50ml) and stirred at r.t. for 30 mins, until TLC (SiO_2; EtOAc:MeOH; 9:1) showed no further reaction. The ethanolic solution of polyallylamine was then added, and the mixture stirred at r.t.

At the same time, FmocLys(Fmoc)-OH (708mg, 1.2mmol), HATU (456mg, 1.2mmol) and DIPEA (418ul, 2.4mmol) were dissolved in DCM (50ml) and stirred for 30 mins. This solution was then added to the reaction containing polyallylamine. The reaction was stirred for 48 hours. Methanol (100ml) was added to the solution and a white precipitate formed. The solid was collected by filtration, and washed with DCM (30ml), and MeOH (50ml). The solid was dried under to vacuum to yield the desired polymer **213** (658mg); **^{1}H NMR** (500MHz, $CDCl_3$ – see appendix 1 for spectrum) δ_H/ppm 1.20-2.00 (br, C$\underline{H}_2$ of polyallylamine and amino acids), 2.60-3.20 (br, C$\underline{H}$ of polyallylamine, C$\underline{H}_2$N of amino acids), 4.70-4.40 (br, $\alpha\underline{H}$ of amino acids), 6.50-7.70 (br, H of Fmoc and Trityl groups); ν_{max} (solid state/cm^{-1}) 732(m), 810(w), 1023(m), 1090(s), 1149(w), 1257(m), 1301(m), 1386(m), 1555(m), 1656(s, C=O), 2935(m), 3246(m, C(O)NH), 3312(m, C(O)NH).

Cleavage of Acid Labile (Trt) Protecting Groups to Afford 222.

Polymer **213** (640mg) was suspended in a solution of TFA:triisopropylsilane:H_2O (95:2.5:2.5, 15ml) and sonicated at r.t. for 5 hours. The mixture was filtered, and washed with H_2O (2 x 40ml), TFA (1 x 50ml), and H_2O (2 x 50ml), and dried under vacuum to yield the desired polymer **222** as an off-white solid (572mg); **^{1}H NMR** (500MHz, CD_3OD – see appendix 1 for spectrum) δ_H/ppm 1.0-1.9 (br, C$\underline{H}_2$ of polyallylamine/ amino acids), 2.6-3.6 (C$\underline{H}$ of polyallylamine, C$\underline{H}_2$NH of amino acids), 3.7-4.4 (br, $\alpha\underline{H}$ of amino acids), 6.50-7.80 (br, Fmoc$\underline{H}$).

Cleavage of Base Labile (Fmoc) Protecting Groups to Afford 216.



216

Polymer **222** (560mg) was suspended in a 20% piperidine/DMF solution (40ml) and sonicated at r.t. for 3 hours. The mixture was filtered and washed with DMF (3 x 40ml), H_2O (2 x 40ml) and DCM (2 x 40ml). The solid was dried under vacuum to yield the fully deprotected polymer **216** as a cream coloured solid (93mg); **^{1}H NMR** (500MHz, D_2O – see appendix 1 for spectra) δ_H/ppm 0.8-1.9 (br, C$\underline{H}_2$ of polyallylamine and amino acids), 2.3-2.4 (br, C$\underline{H}$ of polyallylamine, C$\underline{H}_2$ of cysteine, C$\underline{H}_2$NH of lysine), 3.7-4.4 (br, $\alpha\underline{H}$ of amino acids).

6. Binding and Reactions of Polymers.

6.0 Binding Studies of Polymer 218 using PFG-NMR.

Measurement of the Diffusion Coefficient of Polymer 218.

Polymer **218** (8.5mg) was dissolved in D_2O (1ml) using sonication to ensure complete dissolution. The pD was adjusted to 7 using DCl and NaOD. The diffusion coefficient of the polymer in solution was measured by PFG-NMR using the BPP-LED pulse sequence[125]. A series of 32 BPP-LED spectra were acquired with gradient strength varying on a "squared" scale from 5-95% (100%=50Gcm^{-1}). The parameters used were: δ=6ms, τ=01ms, Δ=200ms. The data was processed as described in section 3. The results are shown in Chapter 2.8.

Measurement of the Diffusion Coefficient of the TSA 146 in a 2mM solution.

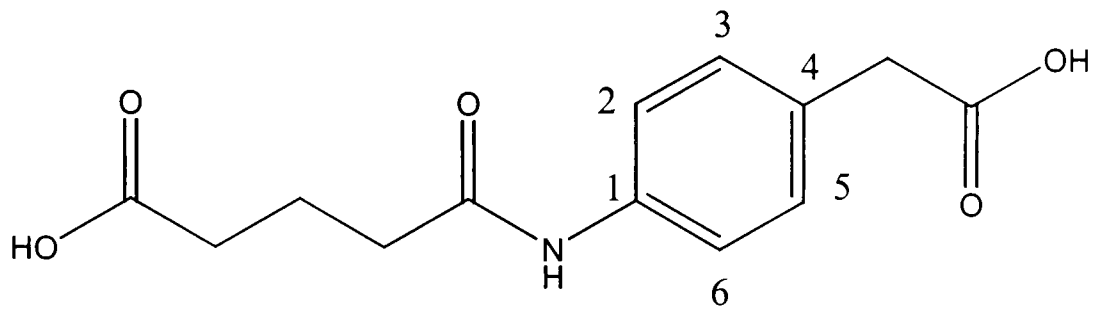
A 2mM solution of TSA **146** in D_2O was prepared, and the pD adjusted to 7 using DCl and NaOD. The diffusion coefficient was measured using the method and parameters described in section 3. The results are shown in Chapter 2.8

Measurement of the Diffusion Coefficient of the TSA 146 in the presence of Polymer 218.

TSA **146** (0.8mg) was added to the sample used to measure the diffusion coefficient of polymer **218**, so that the concentration of TSA **146** in the solution was 2mM. The diffusion coefficient of the TSA **146** was then measured using the same method as described above. The parameters used were: δ=3.2ms, τ=0.1ms, Δ=200ms. The data was processed as described in section 3. The results are shown in Chapter 2.8

6.1 Synthesis of Acid Product 219 (for use as a standard in HPLC).

Synthesis of 4-[(4-Carboxy-1-Oxobutyl)amino] Benzene Acetic Acid (219)



219

4-Aminophenylacetic acid (1g, 6.6mmol) was suspended in 50ml DCM. Glutaric anhydride (0.91g, 7.9mmol) and triethylamine (2.05ml, 14.5mmol) were added and the mixture heated to reflux. After 12 hours a yellow solution had formed. The solution was cooled, and diluted with H_2O (100ml). The organic layer was separated. The pH of the aqueous layer was adjusted to 1 using aq. HCl (2M), and extracted with EtOAc (2 x 100ml). The organic layer was dried over Na_2SO_4, and the solvent removed *in vacuo* to yield **219** as a cream solid (1.65g, 95%); **m.p.** 168-169°C; **^{1}H NMR** (300MHz, D_2O) δ_H/ppm 1.92 (m, 2H, $CH_2C\underline{H}_2CH_2$), 2.30 (t, *J*=7.5, 2H, $C\underline{H}_2C(O)NH$), 2.42 (t, *J*=7.5, 2H, $CH_2C\underline{H}_2CO_2H$), 3.51 (s, 2H, $ArC\underline{H}_2CO_2H$), 7.27 (d, *J*=8.0, 2H, C(3)$\underline{H}$, C(5)$\underline{H}$), 7.36 (d, *J*=8.0, 2H, C(2)$\underline{H}$, C(6)$\underline{H}$); **^{13}C NMR** (D_2O, 75.4MHz) δ_C/ppm 22.7 ($CH_2\underline{C}H_2CH_2$), 36.4 ($\underline{C}H_2C(O)NH$), 37.1 ($CH_2\underline{C}H_2CO_2H$), 44.3 ($Ar\underline{C}H_2CO_2H$), 122.7 ($\underline{C}$(3)H, $\underline{C}$(5)H), 130.1 ($\underline{C}$(2)H, $\underline{C}$(6)H), 134.9 ($\underline{C}$(4)H), 135.4 ($\underline{C}$(1)H), 175.6 ($CH_2\underline{C}(O)NH$), 181.3, 183.0 (2 x $\underline{C}O_2H$); ν_{max} (nujol/cm^{-1}) 653(w), 721(w), 798(m), 895(w, br), 1182(w), 1247(m), 1306(m), 1377(s), 1397(m), 1456(s), 1462(s), 1537(s), 1599(m), 1633(m), 1660(s, C(O)NH), 1699(s, C=O), 1728(s, C=O), 2627(m, br), 2852(s), 2922(s, O-H st), 3300(s, C(O)NH); **m/z** (FAB) 288 (74%, $[M+Na]^+$), 266 (80%, $[M+H]^+$), 248 (65%, $[M+H]^+$-H_2O), 220 (43%, $[M-CO_2H]^+$); **HRMS** found 266.1033, M+H ($C_{13}H_{16}NO_5$) requires 266.1028; **Anal.** Calcd. For $C_{13}H_{15}NO_5$: C, 58.86; H, 5.70; N, 5.28%, Found: C, 58.73; H, 5.71; N, 5.25%.

6.2 Calibration of HPLC by the External Standard Method[165].

The compounds used as standards were all synthesised as described earlier (section 2.0, 2.1). HPLC conditions were: Reverse phase C18 column (4.6 x 250mm), sample injection volume of 2μl, elution of CH_3CN (+5%H_2O, 0.1%TFA)/H_2O (0.1% TFA), solvent gradient 5-95% CH_3CN over 20mins, flow rate 1ml/min, UV detection at 254nM.

6.2.1 Calibration of Ester Starting Material.

A 5mM solution of the ester substrate **145** (46.1mg in 25ml) in aq. phosphate buffer pH7 was prepared. From this stock solution, 5 solutions of differing concentrations (see Table 11) were prepared, and analysed by HPLC. The area of the peak at retention time 18.01 mins. was measured.

Concentration Of Ester	Area of Ester Peak
1mM	606353
0.8mM	506804
0.6mM	367242
0.4mM	240582
0.2mM	127880

Table 11

A plot of the area of the peak against the concentration was made, and linear regression analysis performed (Equation 13):

$y = 615796x$ **Equation 13**

where y is the area of the ester, and x is the concentration.

6.2.2 Calibration of Acid Product.

A 5mM solution of the acid product **219** (33.1mg in 25ml) in aq. phosphate buffer pH7 was prepared. From this stock solution, 5 solutions of differing concentrations (see

Table 12) were prepared, and analysed by HPLC. The area of the peak at retention time 11.9 mins. was measured.

Concentration of Acid	Area of Acid Peak
1mM	556559
0.8mM	437926
0.6mM	331037
0.4mM	216663
0.2mM	111659

Table 12

A plot of the area of the peak against the concentration was made, and linear regression analysis performed (Equation 14):

$y = 551939x$ **Equation 14**

where y is the area of the acid, and x is the concentration.

6.2.3 Calibration of TSA 146.

A 5mM solution of the TSA **146** (50.0mg in 25ml) in aq. phosphate buffer pH7 was prepared. From this stock solution, 4 solutions of differing concentrations (see Table 13) were prepared, and analysed by HPLC. The area of the peak at retention time 13.9 mins. was measured.

Concentration Of TSA	Area of TSA Peak
1mM	587627
0.8mM	479769
0.4mM	237143
0.2mM	122796

Table 13

A plot of the area of the peak against the concentration was made, and linear regression analysis performed (Equation 15):

$y = 592858x$ **Equation 15**

where y is the area of the TSA, and x is the concentration.

6.3 Hydrolysis of Ester 145 with Polymers.

Control Reaction.

1ml of a 0.5mM solution of ester **145** in aq. phosphate buffer pH7 was stirred at r.t. and pressure, and the change in concentration of ester **145**, and formation of acid product **219** were monitored by HPLC at time t=0, t=10mins, and then every 30mins for 7 hours, and then after 24 hours. The results are described in Chapter 2.8

General Procedure.

1ml of a 0.5mM solution of ester **145** in aq. phosphate buffer pH7 was added to 15mg of polymer (1 of **204, 215, 217 or 218**). The mixture was stirred at r.t. and pressure, and the change in concentration of ester **145**, and formation of acid product **219** were monitored by HPLC at time t=0, t=10mins, and then every 30mins for 7 hours, and then after 24 hours. The results are described in Chapter 2.8.

6.4 Inhibition Experiments Using TSA 146.

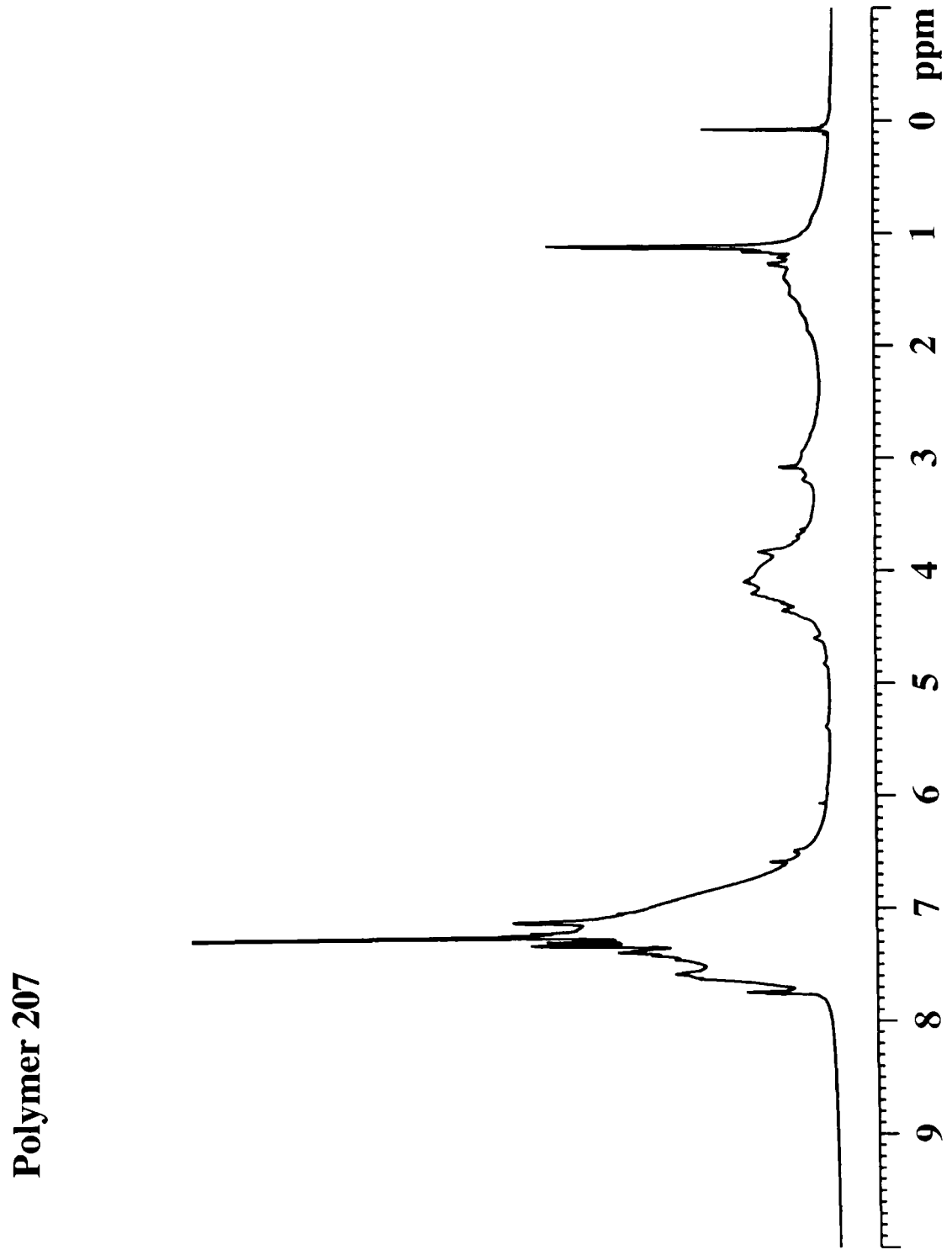
General Procedure.

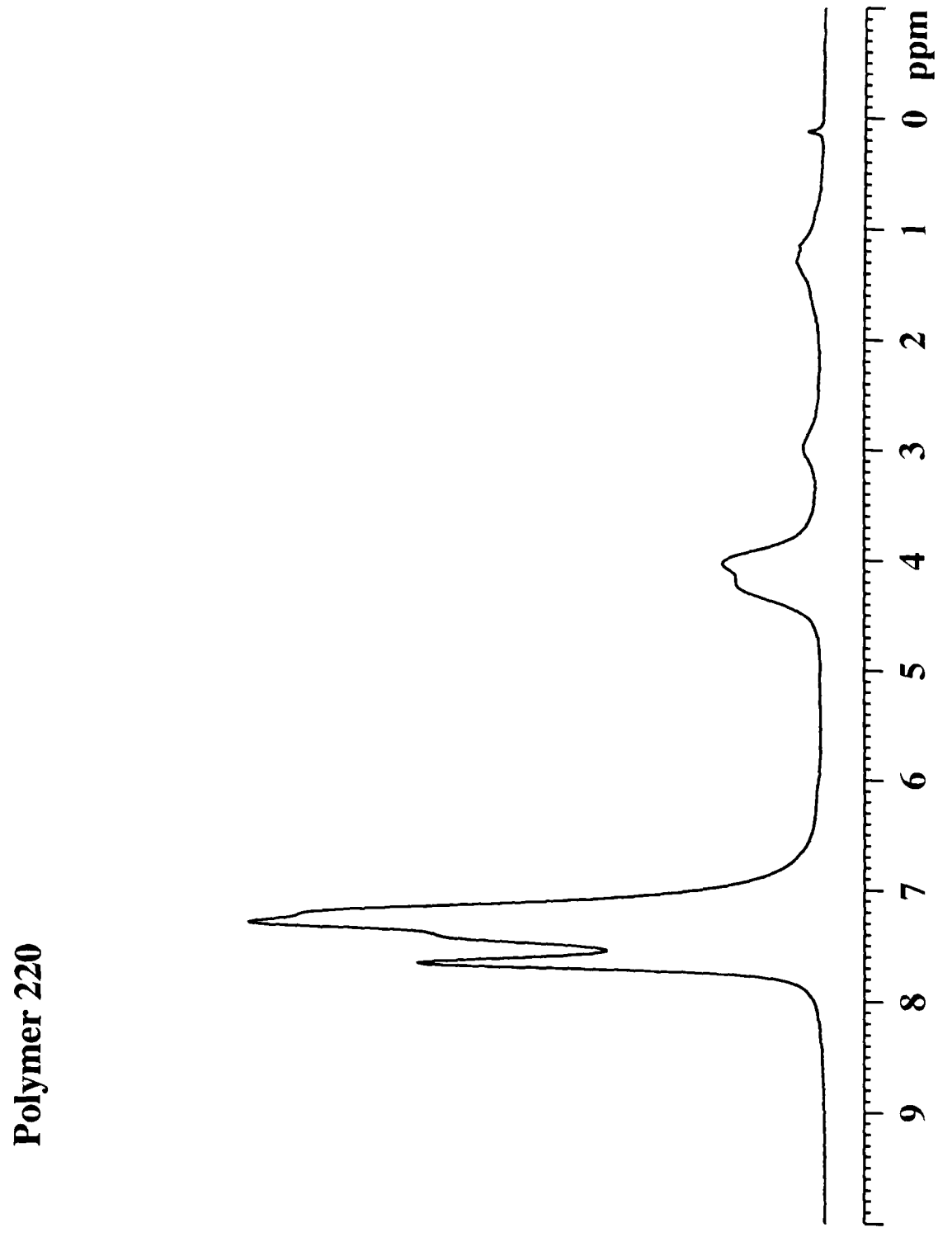
1ml of a solution of concentration 0.5mM ester **145** and 0.5mM TSA **146** in aq. phosphate buffer pH7 was added to 15mg of polymer (1 of **204, 215, 217 or 218**). The mixture was stirred at r.t. and pressure, and and the change in concentration of ester **145**, TSA **146**, and formation of acid product **219** were monitored by HPLC at time t=0, t=10mins, and then every 30mins for 7 hours, and then after 24 hours. The results are described in Chapter 2.8.

Appendix 1.

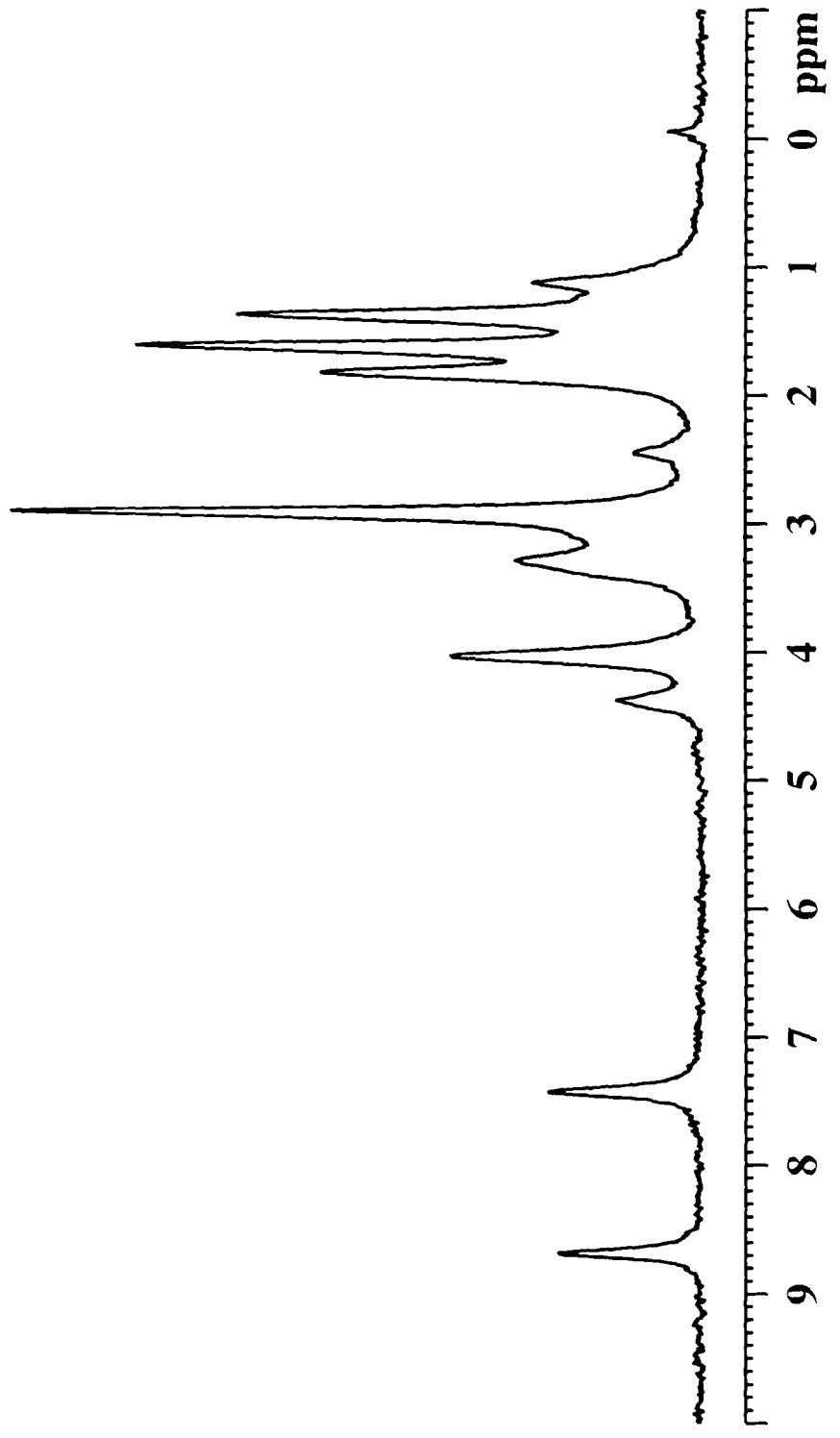
Appendix 1. NMR Spectra of Polymers.

a) Polymers Containing Lysine and Histidine.

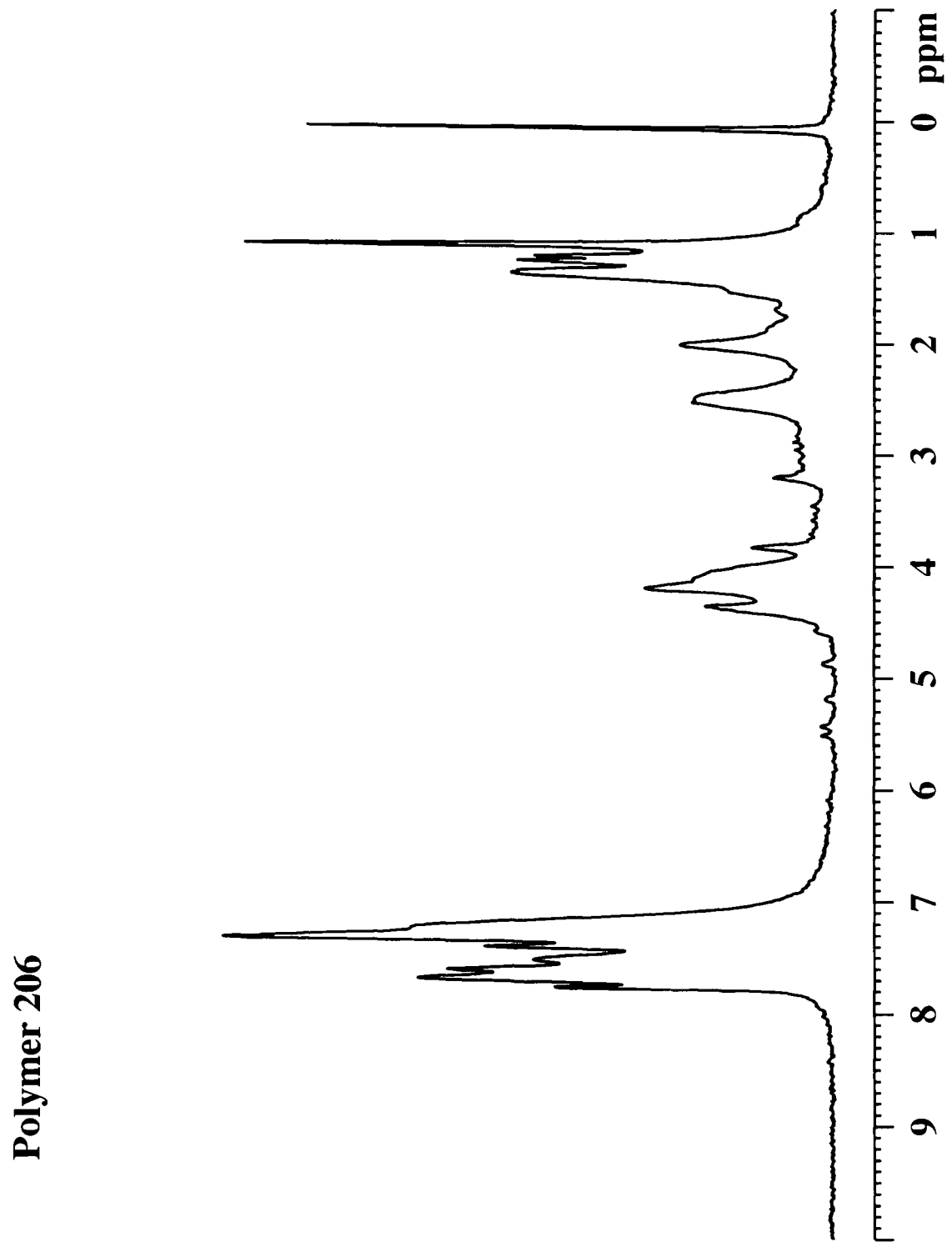


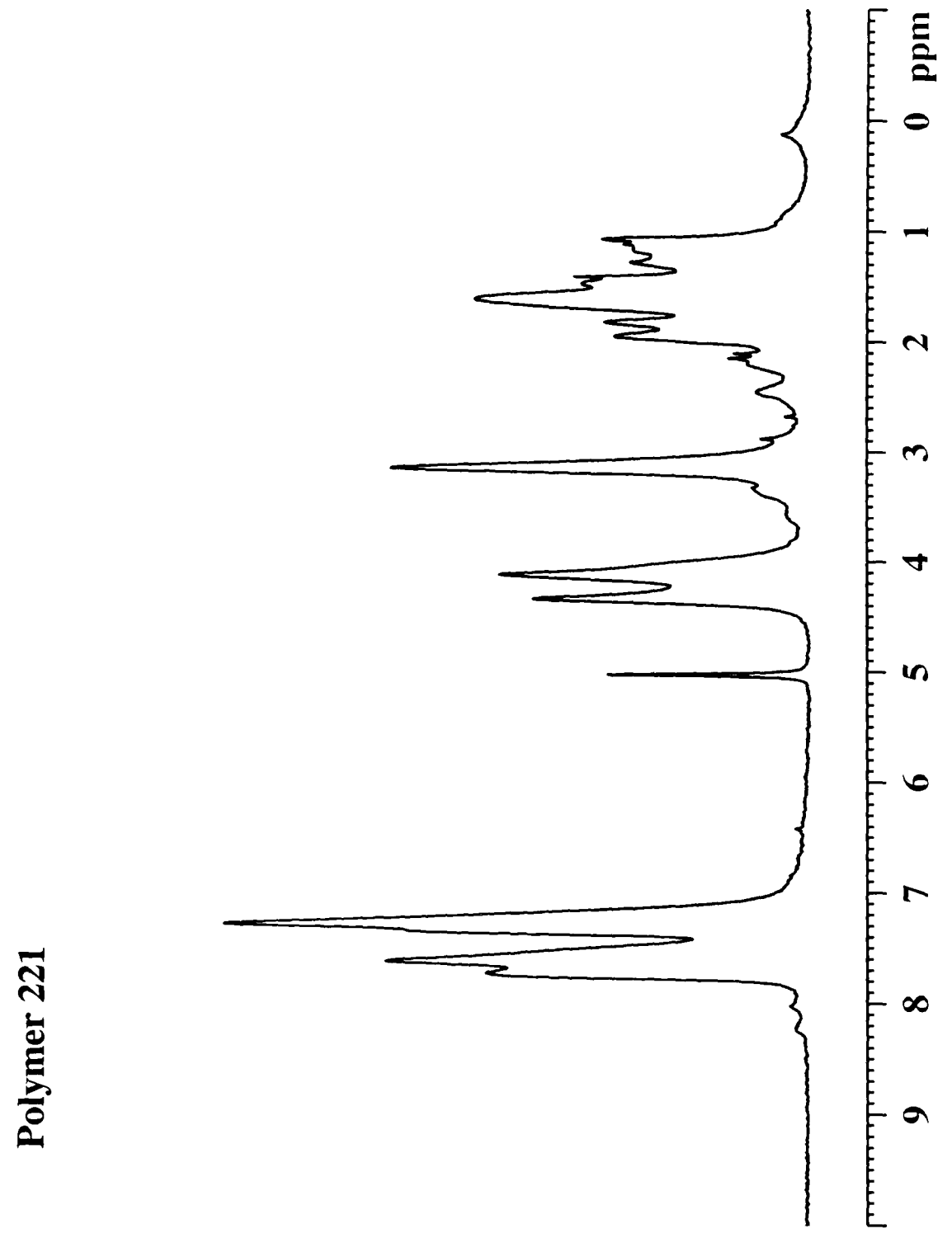

Polymer 220
9
8
7
6
5
4
3
2
1
0 ppm

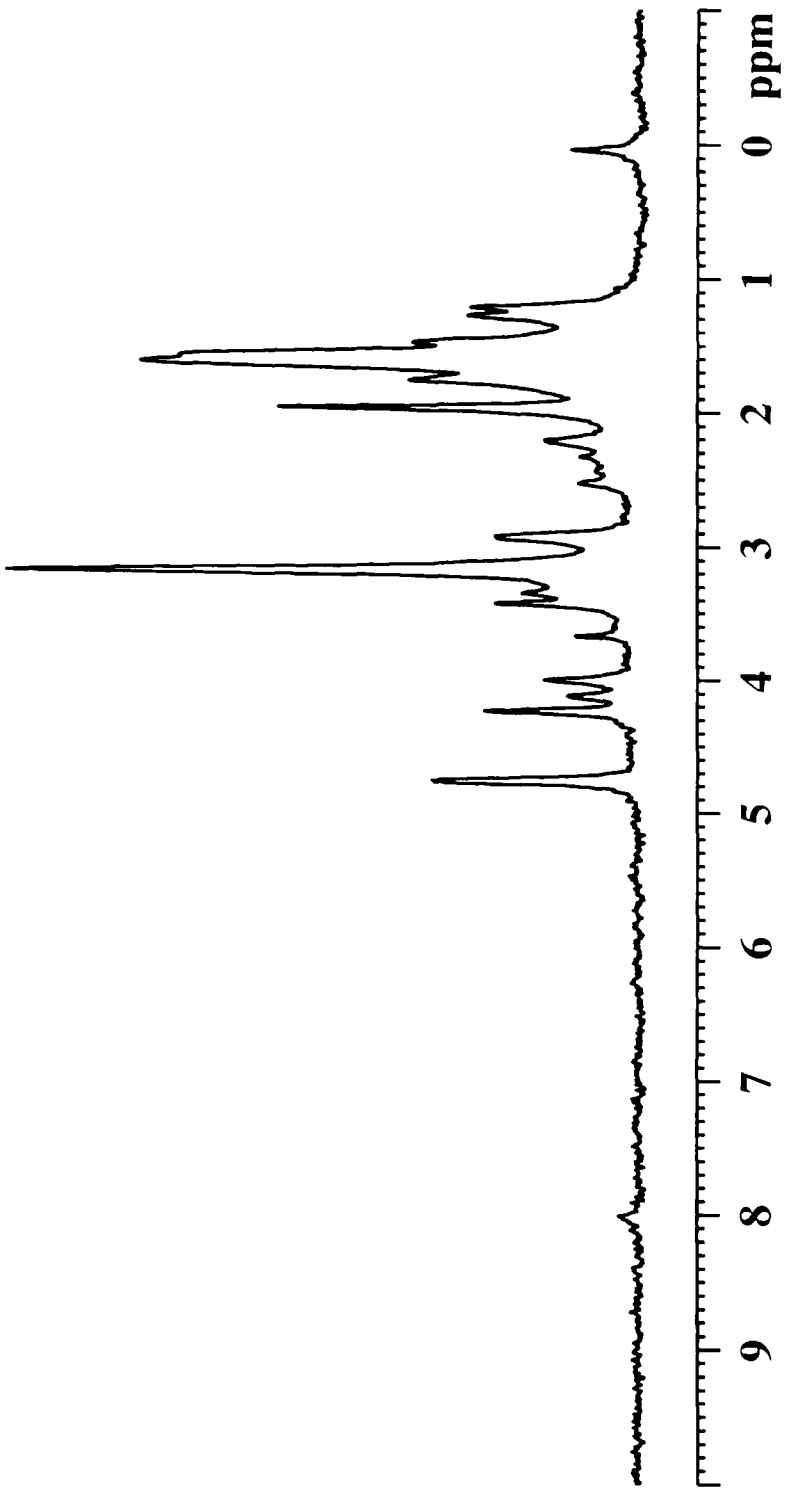
Polymer 217



b) Polymers containing Lysine and Arg-Arg.

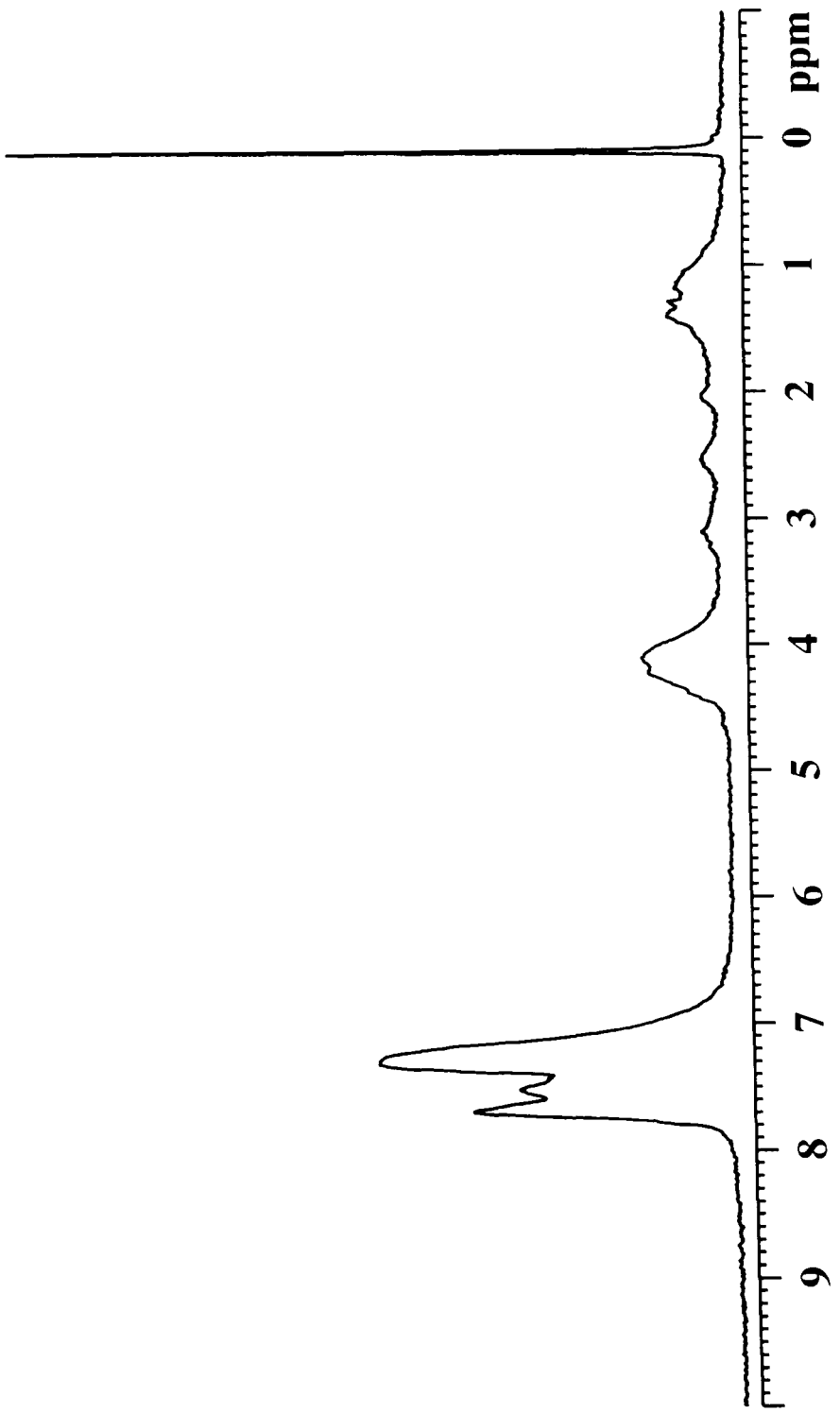



Polymer 221
9
8
7
6
5
4
3
2
1
0 ppm

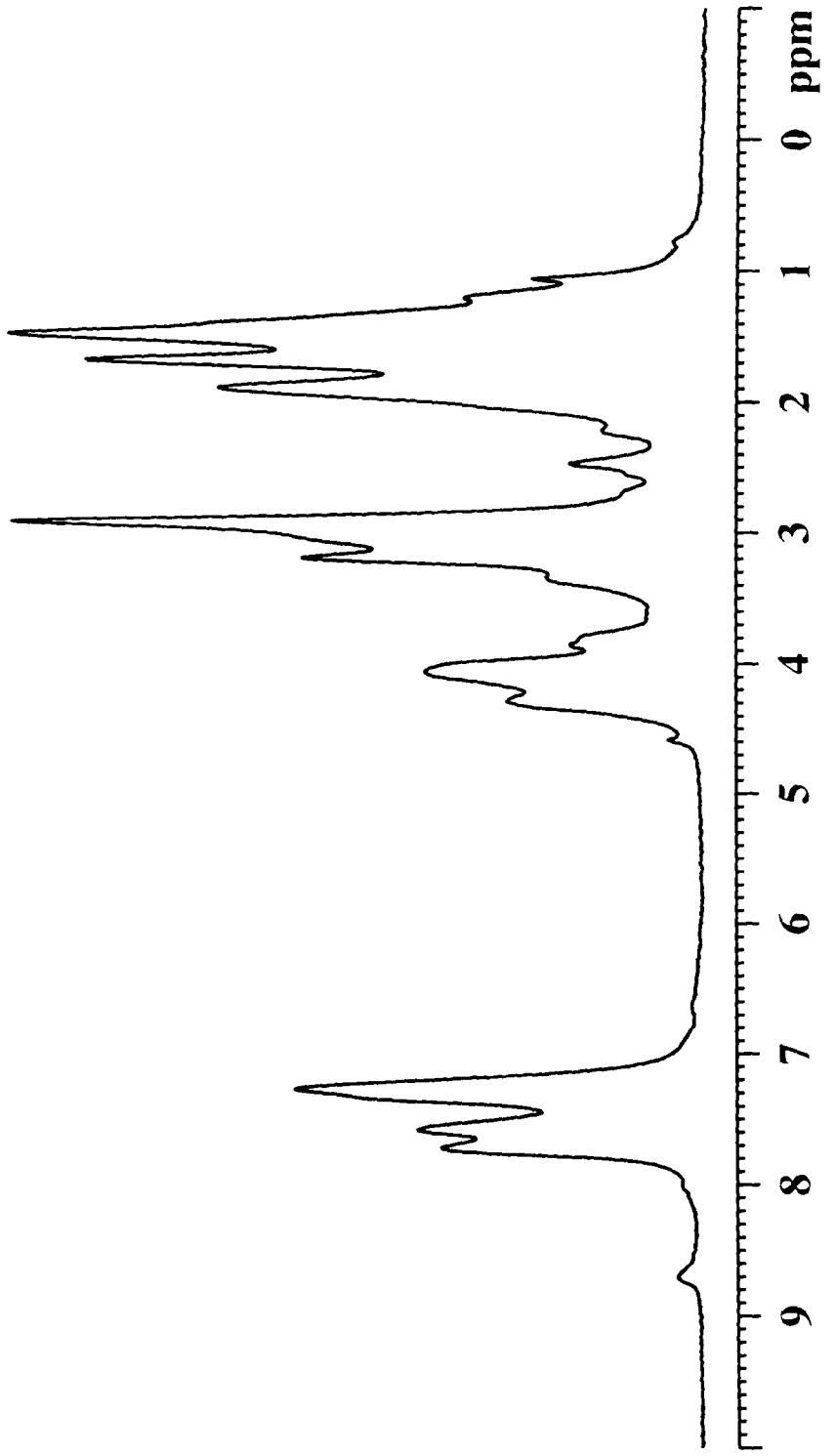


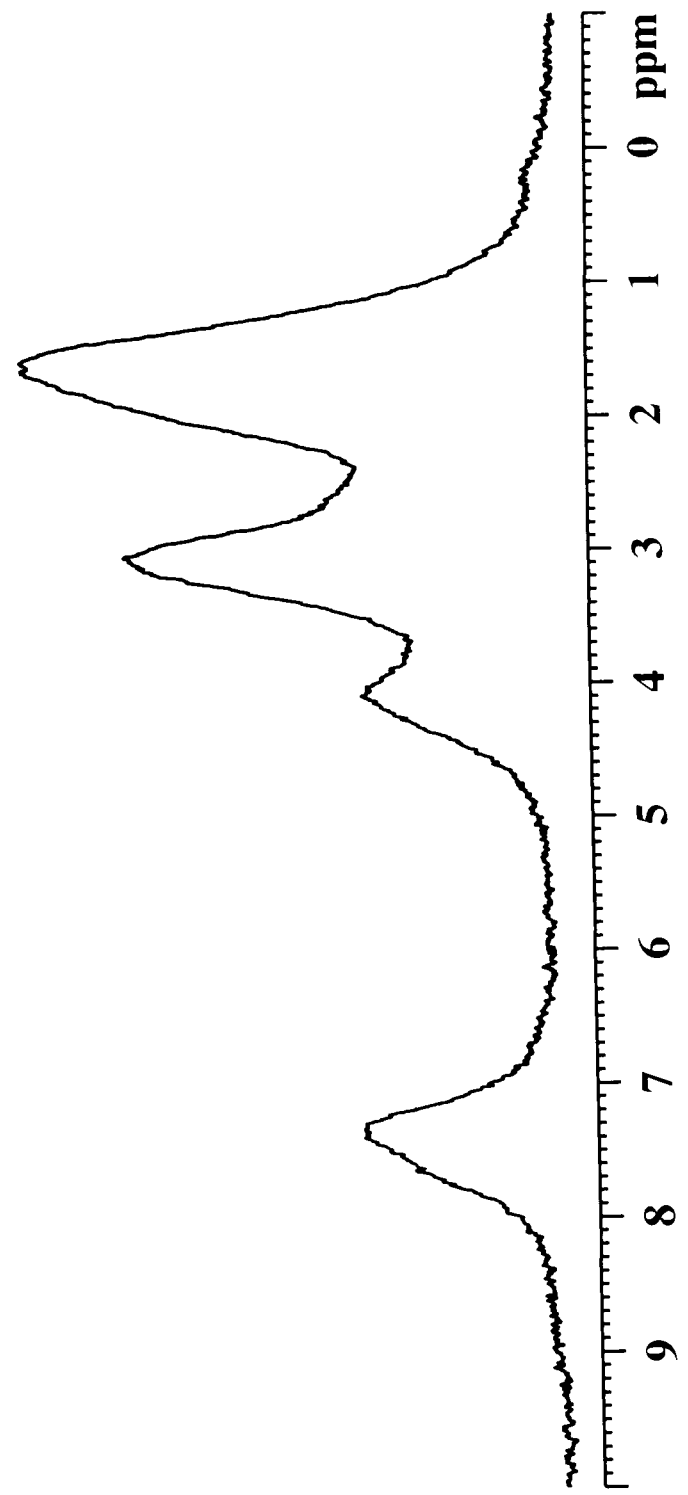
Polymer 218

c) Polymers Containing Lysine, Histidine and Arg-Arg.



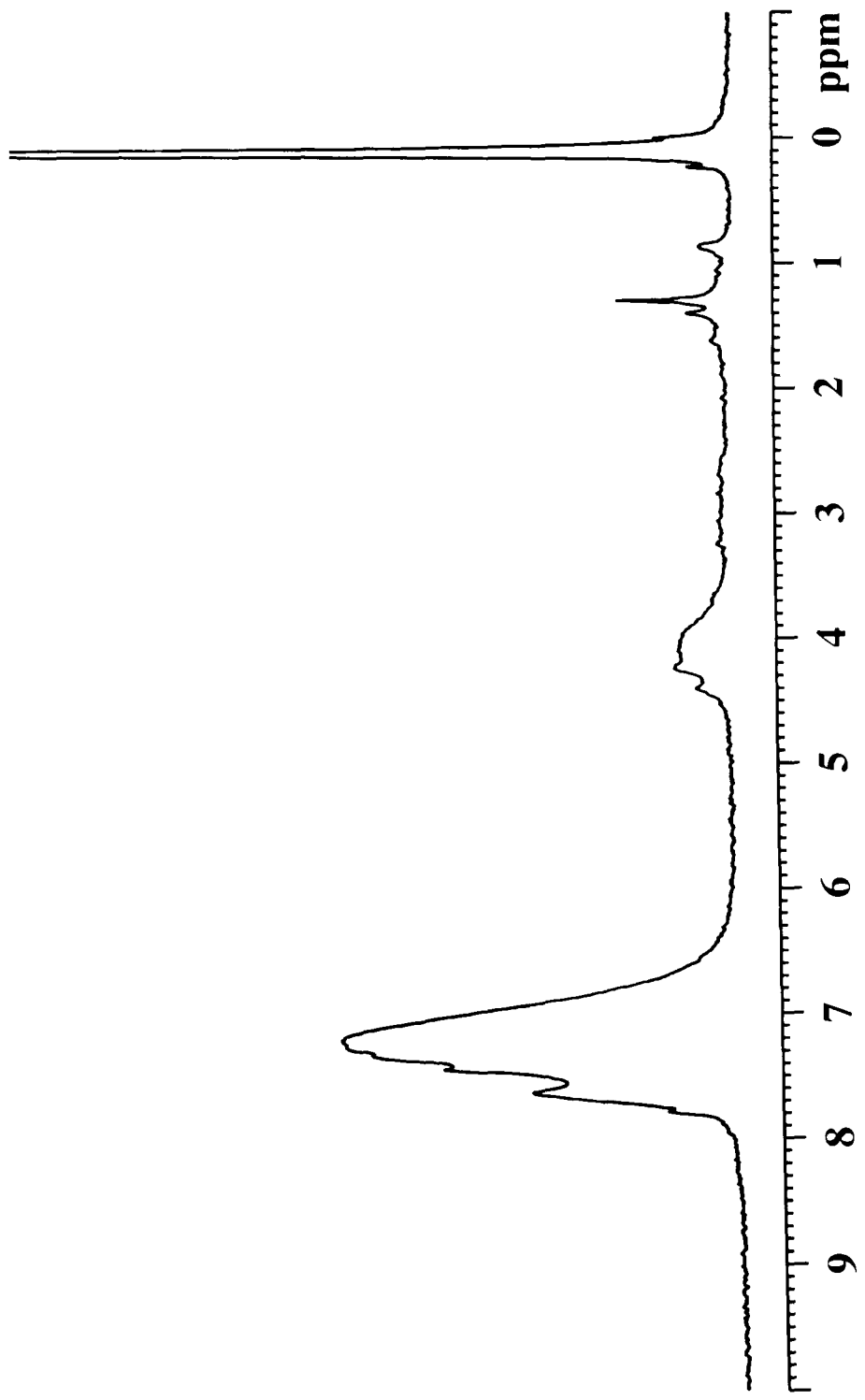
Polymer 208

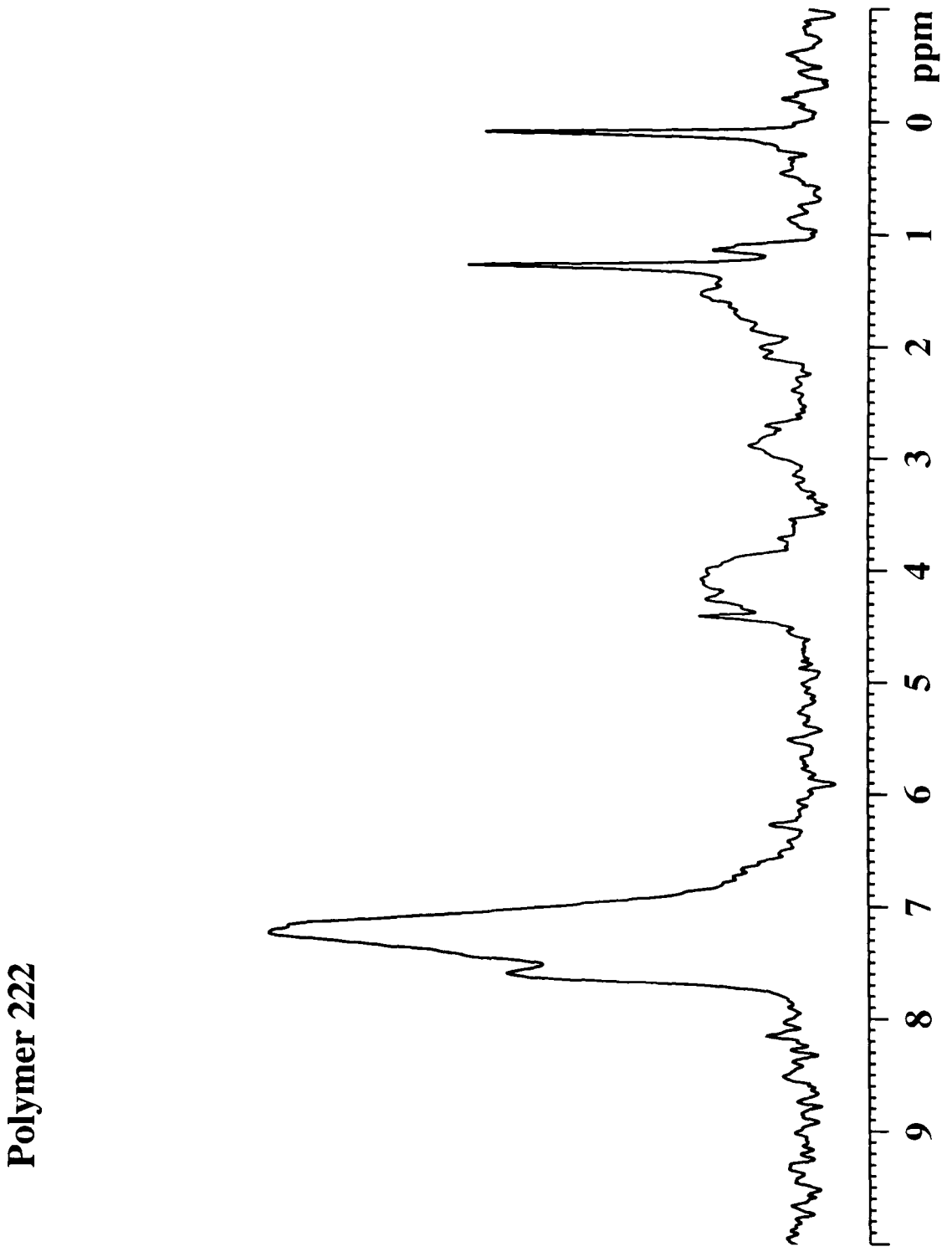

Polymer 214
9
8
7
6
5
4
3
2
1
0 ppm

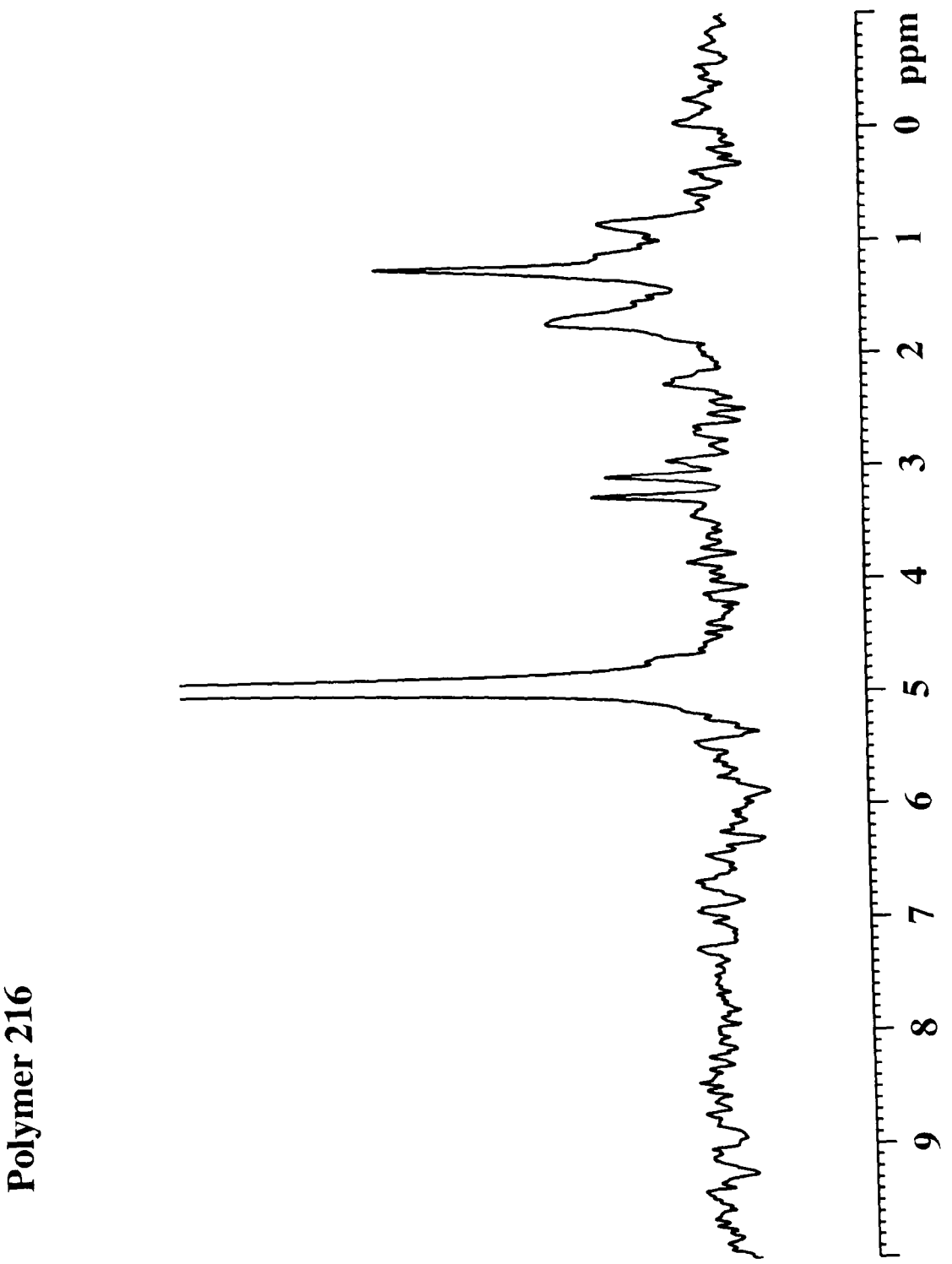


Polymer 215

d) Polymers Containing Lysine and Cysteine.




Polymer 222
9
8
7
6
5
4
3
2
1
0 ppm


Polymer 216
0 ppm
1
2
3
4
5
6
7
8
9

References.

References:

1. Fischer, E., *Ber. Dtsch. Chem. Ges.*, **1894**, *27*, 2985.

2. Fersht, A. R., 'Enzyme Structure and Mechanism, 2nd Ed'., Freeman New York, **1985**.

3. Pauling, L., *Nature*, **1948**, *161*, 707.

4. Rowan, S. J.; Sanders, J. K. M., *Curr. Opin. Chem. Biol.*, **1997**, *1*, 438.

5. Blow, D. M., *Acc. Chem. Res.*, **1976**, *9*, 146.

6. Blow, D. M.; Birktoft, J. J.; Hartley, B. S., *Nature*, **1969**, *221*, 337.

7. Blow, D. M.; Sigler, P. B.; Matthew, B. W.; Henderson, R., *J. Mol. Biol.*, **1968**, *35*, 143.

8. Mader, M. A.; Bartlett, P. A., *Chem. Rev.*, **1997**, *97*, 1281.

9. Eriksson, A. E.; Baase, W. A.; Wozniak, J. A.; Matthew, B. W., *Nature*, **1992**, *355*, 371.

10. Davis, A. M.; Teague, S. J., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 707.

11. Williams, D. H.; Searle, M. S.; Mackay, J. P.; Gerhard, V.; Maplestone, R. A., *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 1172.

12. Fersht, A. R.; Shi, J.-P.; Knill-Jones, J.; Lowe, D. M.; Wilkinson, A. J.; Blow, D. M.; Brick, P.; Carter, P.; Waye, M. M. Y.; Winter, G., *Nature*, **1985**, *314*, 235.

13. Hansch, C., *Acc. Chem. Res.*, **1993**, *26*, 147.

14. Hansch, C.; Hoekman, D.; Gao, H., *Chem. Rev.*, **1996**, *96*, 1045.

15. Tong, L.; Pov, S.; Mui, S.; Lamarre, D.; Yockim, C.; Beaulieu, P.; Anderson, P. C., *Structure*, **1995**, *3*, 33.

16. Kirby, A., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 707.

17. Komiyana, M.; Inoue, S., *Bull. Chem. Soc. Jpn.*, **1980**, *52*, 334.

18. Komiyana, M.; Hirai, H., *Chem. Lett.*, **1980**, 1251.

19. Desper, J. M.; Breslow, R., *J. Am. Chem. Soc.*, **1994**, *116*, 12081.

20. Breslow, R.; Nesnas, N., *Tetrahedron Lett.*, **1999**, *40*, 3335.

21. Tastan, P.; Akkaya, E. U., *J. Mol. Catal. A: Chem.*, **2000**, *157*, 261.

22. Yu, J.; Zhao, Y.; Hofferman, M. J.; Venton, D. L., *Bioorg. Med. Chem. Lett.*, **1999**, *9*, 2705.

23. Mattei, P.; Diederich, F., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 1341.

24. Häring, D.; Distefano, M. D., *Bioconjugate Chem.*, **2001**, *12*, 385.

25. Breslow, R.; Dong, S. D., *Chem. Rev.*, **1998**, *98*, 1997.

26. Breslow, R.; Zhang, B., *J. Am. Chem. Soc.*, **1992**, *114*, 5882.

27. Clyde-Watson, Z.; Vidal-Ferron, A.; Twyman, L. J.; Walker, C. J.; McCallien, W. J.; Fanni, S.; Bampos, N.; Wylie, R. S.; Sanders, J. K. M., *New. J. Chem.*, **1998**, 493.

28. Marty, M.; Clyde-Watson, Z.; Twyman, L. J.; Nakash, M.; Sanders, J. K. M., *J. Chem. Soc. Chem. Commun.*, **1998**, 2265.

29. Sanders, J. K. M., *Pure Appl. Chem.*, **2000**, *72*, 2265.

30. Laschat, S., *Angew. Chem. Int. Ed. Engl.*, **1996**, *35*, 289.

31. Kang, J.; Santamaria, J.; Hilmersson, G.; Rebek Jr., J., *J. Am. Chem. Soc.*, **1998**, *120*, 7389.

32. Meissner, R. S.; Rebek Jr., J.; de Mendoza, J., *Science*, **1995**, *270*, 1485.

33. Kang, J.; Rebek Jr., J., *Nature*, **1997**, *385*, 50.

34. Kang, J.; Hilmerson, G.; Santamaria, J.; Rebek Jr., J., *J. Am. Chem. Soc.*, **1998**, *120*, 3650.

35. Tam-Chang, S.-W.; Jimenez, L.; Diederich, F., *Helv. Chim. Acta.*, **1993**, *76*, 2616.

36. Mattei, P.; Diederich, F., *Helv. Chim. Acta.*, **1997**, *80*, 1555.

37. Habicher, T.; Diederich, F.; Gramlich, B., *Helv. Chim. Acta.*, **1999**, *82*, 1066.

38. Klotz, K. I.; Royer, G. P.; Scarpa, I. S., *Proc. Natl. Acad. Sci. USA.*, **1971**, *68*, 263.

39. Jang, B-B.; Lee, K.-P.; Min, D.-H.; Suh, J., *J. Am. Chem. Soc.*, **1998**, *120*, 12008.

40. Suh, J.; Moon, S.-J., *Inorg. Chem.*, **2001**, *40*, 4890.

41. Jeung, C.-S.; Kim, C. H.; Min, K.; Suh, S.-W.; Suh, J., *Bioorg. Med. Chem. Lett.*, **2001**, *11*, 2401.

42. Moon, S.-J.; Jeon, J. W.; Kim, H.; Suh, M. P.; Suh, J., *J. Am. Chem. Soc.*, **2000**, *122*, 7742.

43. Suckling, C. J.; Zhu, L. M., *Bioorg. Med. Chem. Lett.*, **1993**, *3*, 531.

44. Häring, D.; Schuler, E.; Adam, W.; Saha-Moller, C. R.; Schreier, P., *J. Org. Chem.*, **1999**, *64*, 832.

45. Goldberg, J. M.; Kirsch, J. F., *Biochemistry*, **1996**, *35*, 5280.

46. Pauling, L., *Chem. Eng. New.*, **1946**, *24*, 1375.

47. Jencks, W. P., 'Catalysis in Chemistry and Enzymology', McGraw-Hill, New York, **1969**.

48. Raso, V.; Stollar, B. D., *Biochemistry*, **1975**, *14*, 591.

49. Kohler, G.; Milstein, C., *Nature*, **1975**, *256*, 495.

50. Tramontano, A.; Janda, K. D.; Lerner, R. A., *Science*, **1986**, *234*, 1566.

51. Pollack., S. J.; Jacobs, J. W.; Schultz, P. G., *Science*, **1986**, *234*, 1570.

52. Benkovic, S. J.; Adams, J. A.; Bordes Jr., C. L.; Janda, K. D.; Lerner, R. A., *Science*, **1991**, *252*, 680.

53. Hsieh, L. C.; Yonkovich, S.; Kochersperger, L.; Schultz, P. G., *Science*, **1993**, *260*, 337.

54. Hilvert, D.; Hill, K. W.; Nared, K. D.; Auditor, M. T. M., *J. Am. Chem. Soc.*, **1989**, *111*, 9261.

55. Gouverneur, V. E.; Houk, K. N.; de Pascual-Theresa, B.; Beno, B.; Janda, K. D.; Lerner, R. A., *Science*, **1993**, *262*, 204.

56. Gruber, K.; Zhour, B.; Houk, K. N.; Lerner, R. A.; Shevlin, C. G.; Wilson, I. A., *Biochemistry*, **1999**, *38*, 7062.

57. Hsieh, L. C.; Schultz, P. G.; Stephans, J. C., *J. Am. Chem. Soc.*, **1994**, *116*, 2167.

58. Fujii, I.; Fukuyama, S.; Iwabuchi, Y.; Tanimura, R., *Nat. Biotechnol.*, **1998**, *16*, 463.

59. Driggers, E. M.; Cho, H. S.; Lui, C. W.; Katzka, C. P.; Braisted, A. C.; Ulrich, H. D.; Wenner, D. E.; Schultz, P. G., *J. Am. Chem. Soc.*, **1998**, *120*, 1995.

60. Li, T.; Janda, K. D.; Lerner, R. A., *Nature*, **1996**, *379*, 326.

61. Wirsching, P.; Ashley, J. A.; Lo, C.-H.; Janda, K. D.; Lerner, R. A., *Science*, **1995**, *270*, 1775.

62. Hasserodt, J., *Synlett*, **1999**, 2007.

63. Janda, K. D.; Weinhouse, M. I.; Schloeder, D. M.; Lerner, R. A.; Benkovic, S. J., *J. Am. Chem. Soc.*, **1990**, *112*, 1274.

64. Ersoy, O.; Fleck, R.; Sinskey, A.; Masamune, S., *J. Am. Chem. Soc.*, **1996**, *118*, 13077.

65. Zhong, G.; Lerner, R. A.; Barbas III, C. F., *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 3738.

66. Wagner, J.; Lerner, R. A.; Barbas III, C. F., *Science*, **1997**, *270*, 1797.

67. Zhong, G.; Shabat, D.; List, B.; Anderson, J.; Sinha, S. C.; Lerner, R. A.; Barbas III, C. F., *Angew. Chem. Int. Ed. Engl.*, **1998**, *110*, 2609.

68. Abe. I.; Rohmer, M.; Prestwich, G. D., *Chem. Rev.*, **1993**, *93*, 2189.

69. Paschall, C. M.; Hasserodt, J.; Jones, T.; Lerner, R. A.; Janda, K. D.; Christianson, D. W., *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 1743.

70. Hasserodt, J.; Janda, K. D.; Lerner, R. A., *Bioorg. Med. Chem.*, **2000**, *8*, 995.

71. Wulff, G., *Angew. Chem. Int. Ed. Engl.*, **1995**, *34*, 1812.

72. Vulfson, E.; Alexander, C.; Whitcombe, M., *Chem. Brit.*, **1997**, 23.

73. Wulff, G.; Gross, T.; Schonfeld, R., *Angew. Chem. Int. Ed. Engl.*, **1997**, *36*, 1962.

74. Matsui, J.; Nicholls, I. A.; Korube, I.; Mosbach, K., *J. Org. Chem.*, **1996**, *61*, 5414.

75. Liu, X.-C.; Mosbach, K., *Macromol. Rapid Commun.*, **1997**, *18*, 609.

76. Strikovsky, A. G.; Kasper, D.; Grun, M.; Green, B. S.; Hradil, J.; Wulff, G., *J. Am. Chem. Soc.*, **2000**, *122*, 6295.

77. Sellergren, B.; Karmalkar, R. N.; Shea, K. J., *J. Org. Chem.*, **2000**, *65*, 4009.

78. Polborn, K.; Severin, K., *J. Chem. Soc. Chem. Commun.*, **1999**, 2481.

79. Polborn, K.; Severin, K., *Chem. Eur. J.*, **2000**, *6*, 4604.

80. Suh, J.; Hah, S. S., *J. Am. Chem. Soc.*, **2000**, *122*, 3901.

81. Oh, S.; Chang, W.; Suh, J., *Bioorg. Med. Chem. Lett.*, **2001**, *11*, 1469.

82. Klibanov, A. M., *Trends. Biochem. Sci.*, **1989**, *14*, 141.

83. Klibanov, A. M., *Nature*, **1995**, *374*, 536.

84. Braco, L.; Dabulis, K.; Klibanov, A. M., *Proc. Natl. Acad. Sci. USA.*, **1990**, *87*, 274.

85. Slade, C. J.; Vulfson, E. N., *Biotech. Bioeng.*, **1998**, *57*, 212.

86. Peiβker, F.; Fischer, L., *Bioorg, Med. Chem.*, **1999**, *7*, 2231.

87. Huc, I.; Lehn, J.-M., *Proc. Natl. Acad. Sci. USA.*, **1997**, *94*, 2106.

88. Eliseev, A. V.; Nelen, M. I., *Chem. Eur. J.*, **1998**, *4*, 825.

89. Furlan, R. L. E.; Cousins, G. R. L.; Sanders, J. K. M., *J. Chem. Soc. Chem. Commun.*, **2000**, 1761.

90. Huc, I.; Krische, M. J.; Funeriu, D. P.; Lehn, J.-M., *Eur. J. Inorg. Chem.*, **1999**, 1415.

91. Polyakv, V. A.; Nelen, M. I.; Nazapack-Kandlousy, N; Ryabov, A. D.; Eliseev, A. V., *J. Phys. Org. Chem.*, **1999**, *12*, 357.

92. Cousins, G. R. L.; Poulsen, S.-A.; Sanders, J. K. M., *J. Chem. Soc. Chem. Commun.*, **1999**, 1575.

93. Karan, C.; Miller, B. L., *J. Am. Chem. Soc.*, **2001**, *123*, 7455.

94. Goral, V.; Nelen, M. I.; Eliseev, A. V.; Lehn, J.-M., *Proc. Natl. Acad. Sci. USA.*, **2001**, *98*, 1347.

95. Menger, F. M.; Eliseev, A. V.; Migulin, B. A., *J. Org. Chem.*, **1995**, *60*, 6666.

96. Menger, F. M.; West, C. A.; Ding, J., *J. Chem. Soc. Chem. Commun.*, **1997**, 663.

97. Menger, F. M.; Ding, J.; Barragan, V., *J. Org. Chem.*, **1998**, *63*, 7578.

98. Berkessel, A.; Hérault, D. A., *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 102.

99. Jandeleit, B.; Schaefer, D. J.; Powers, T. S.; Turner, H. W.; Weinberg, W. H., *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 2494.

100. Bein. T., *Angew. Chem. Int. Ed. Engl.*, **1999**, *38*, 323.

101. Reetz, M. T., *Angew. Chem. Int. Ed. Engl.*, **2001**, *40*, 284.

102. Sutherland, J. D., *Curr. Opin. Chem. Biol.*, **2000**, *4*, 263.

103. Hart, E. A.; Hua, L.; Darr, L. B.; Wilson, W. K.; Pang, J.; Matsuda, S. P. T., *J. Am. Chem. Soc.*, **1999**, *121*, 9887.

104. Oue, S.; Okamoto, A. Y. T.; Kagamiyama, H., *J. Biol. Chem.*, **1999**, *274*, 2344.

105. Reetz, M. T.; Zonta, A.; Shimossek, K.; Liebeton, K.; Jaeger, K.-E., *Angew. Chem. Int. Ed. Engl.*, **1997**, *36*, 2831.

106. Clayden, J.; Greeves, N.; Warren, S.; Wothers, P., 'Organic Chemistry', Oxford University Press, **2001**.

107. Wolfenden, R., *Annu. Rev. Biophys. Bioeng.*, **1976**, *5*, 271.

108. Lugoukin; Arbuzov, *Doklady Akad. Nauk. S.S.S.R.*, **1948**, *59*, 1309.

109. Kagan, G.; Birkenmayer, R. D.; Strube, R. E., *J. Am. Chem. Soc.*, **1959**, *81*, 3026.

110. Smith, M. B., 'Organic Synthesis', McGraw Hill, New York, **1994**.

111. Moree, W. J.; van der Morel, G. A.; van Boom, J. H.; Liskamp, R. M. J., *Tetrahedron*, **1993**, *49*, 11055.

112. Balthazar, T. M.; Flores, R. A., *J. Org. Chem.*, **1980**, *45*, 529.

113. Malachowski, W. P.; Coward, J. P., *J. Org. Chem.*, **1994**, *45*, 7616.

114. Malachowski, W. P.; Coward, J. P., *J. Org. Chem.*, **1994**, *45*, 7625.

115. Janda, K. D.; Benkovic, S. J.; Lerner, R. A., *Science*, **1989**, *244*, 437.

116. Campbell, D. A., *J. Org. Chem.*, **1992**, *57*, 6331.

117. March, J., 'Advanced Organic Chemistry: Reaction, Mechanisms and Structure., Wiley, New York, **1992**.

118. Nicholls, I., *J. Molecular Recognition*, **1998**, *11*, 79.

119. Eliseev, A. E., *Current Opinion in Drug Discovery and Development*, **1998**, *1*, 106.

120. Roberts, G. C. K., *Drug Discovery Today*, **2000**, *5*, 230.

121. Dowden, J.; Edwards, P. D.; Flack, S. S.; Kilburn, J. D., *Chem. Eur. J.*, **1999**, *5*, 79.

122. Orfi, L.; Lin, M.; Larive, C. K., *Anal. Chem.*, **1998**, *70*, 1339.

123. Frish, L.; Sansone, F.; Casnati, A.; Ungaro, R.; Cohen, Y., *J. Org. Chem.*, **2000**, *65*, 5026.

124. Gounarides, J. S.; Chen, A.; Shapiro, M. J., *J. Chromatogr. B.*, **1999**, *725*, 70.

125. Wu, D.; Chen, A.; Johnson, C. S., *J. Magn. Reson. A.*, **1995**, *115*, 260.

126. Stejskal, E. O.; Tanner, J. E., *J. Chem. Phys.*, **1965**, *42*, 288.

127. Data obtained using ISOSTAR, accessed from the Chemical Database Service at Daresbury, see ref 137.

128. Fielding, L., *Tetrahedron*, **2000**, *56*, 6151.

129. Carcanague, D. R.; Diederich, F., *Angew. Chem. Int. Ed. Engl.*, **1990**, *29*, 769.

130. Peterson, B. R.; Wallimann, P.; Carcanague, D. R.; Diederich, F., *Tetrahedron*, **1995**, *51*, 401.

131. Sybyl 6.6.2 Tripos Inc., 1699 South Hanley Road., St. Louis, Missouri, 63144, USA.

132. Grail, B. M.; Payne, J. W., *J. Pept. Sci.*, **2000**, *6*, 186.

133. Cohen, N. C., 'Guidebook on Molecular Modelling in Drug Design', Academic Press, **1996**.

134. Goodman, J. M., 'Chemical Applications of Molecular Modelling', Royal Society of Chemistry, Cambridge, **1998**.

135. Ghose, R. S.; Jaeger, E. P.; Kowalazyk, P. J.; Peter, M. L.; Treasurywala, A. M., *J. Comput. Chem.*, **1993**, *14*, 1050.

136. Neuhaus D.; Williams, M. P., 'The Nuclear Overhauser Effect in Structural and Conformational Analysis 2nd Ed.', Wiley-VCH, **2000**.

137. Isostar is accessed from the Chemical Database Service at Daresbury.

Fletcher, D. A.; McMeaking, R. F.; Parkin, D. J., *Chem. Inf. Comput. Sci.*, **1996**, *36*, 746.

138. Bruno, I. J.; Cole, J. C.; Lommerse, J. P. M.; Rowland, R. S.; Taylor, R.; Verdonk, M. L., *J. Comput. Aided Mol Des.*, **1997**, *11*, 525.

139. Allen, F. M.; Kennard, O., *Chemical Design Automation New.*, **1993**, *8*, 31.

140. Fersht, A. R., 'Structure and Mechanism in Protein Science: A Guide to Enzyme Catalysis and Protein Folding', W. H. Freeman, New York, **1999**.

141. Carpino, L. A.; Shroff, H.; Triolo, S. A.; Mansour, E.-S. M. E.; Wenschuh, H.; Albericio, F., *Tetrahedron Lett.*, **1993**, *34*, 7829.

142. Wang, S.-S., *J. Am. Chem. Soc.*, **1973**, *95*, 1328.

143. Barlos, K.; Chatzi, O.; Gatos, D.; Stavropoulos, G., *Int. J. Peptide. Protein. Res.*, **1991**, *37*, 513.

144. Carpino, L. A.; Han, G. Y., *J. Org. Chem.*, **1972**, *37*, 3404.

145. Chan, W. C.; White, P. D., 'Fmoc Solid Phase Peptide Synthesis: A Practical Approach', Oxford University Press, **2000**.

146. Lapatsanis, L.; Milias, G.; Froussios, K.; Kolovos, M., *Synthesis*, **1983**, 671.

147. Tanaka, N.; Nakagawa, K.; Iwasaki, H.; Hosoya, K.; Kimata, K.; Araki, T.; Patterson Jr., D. G., *J. Chromatogr. A.*, **1997**, *781*, 139.

148. Sehgal, D.; Vijay, I. K., *Anal. Biochem.*, **1994**, *218*, 87.

149. Sheehan, J. C.; Ledis, S. L., *J. Am. Chem. Soc.*, **1973**, *95*, 875.

150. Pennington, R. M.; Fisher, R. R., *J. Biol. Chem.*, **1981**, *256*, 8963.

151. Angell, Y. M.; Garcia-Echeverna, C.; Rich, D. H., *Tetrahedron Lett.*, **1994**, *34*, 5981.

152. Williams, D. H.; Fleming, I., 'Spectroscopic Methods in Organic Chemistry, 5th Ed.', McGraw Hill, UK, **1995**.

153. Carpino, L. A., *J. Am. Chem. Soc.*, **1993**, *115*, 4397.

154. White, P., *NovaBioChem Innovations*, 2/93.

155. Pearson, D.; Blanchette, M.; Baker, M. L.; Guidon, C. A., *Tetrahedron Lett.*, **1989**, *30*, 2739.

156. Weingartner, *Z. Phys. Chem. (Neue Folge)*, **1982**, 129.

157. Holz, M.; Weingartner, *J. Magn. Reson.*, **1991**, *92*, 115.

158. Thuong, N. T., *Bull. Soc. Chim. Fr.*, **1965**, 1925.

159. Scherer, H., *Chem. Ber.*, **1972**, *105*, 3357.

160. Rewcastle, G. W.; Baguley, B. C.; Cain, B. F.; *J. Med. Chem.*, **1982**, *25*, 1231.

161. Janda, K. D.; Ashley, J. A.; Jones, T. M.; McLeod, D. A.; Schloeder, D. M.; Weinhouse, M. I.; Lerner, R. A.; Gibbs R. A.; Benkovic, P. A.; Hilhorst, R.; Benkovic, S. J., *J. Am. Chem. Soc.*, **1991**, *112*, 291.

162. Tramontano, A.; Ammann, A. A.; Lerner, R. A., *J. Am. Chem. Soc.*, **1988**, *110*, 2282.

163. Berger, S.; Braun, S.; Kalinowski, H.-O., '150 and More Basic NMR Experiments', Wiley-VCH, **1998**.

164. Fujii, N.; Futaki, S.; Funakoshi, S.; Akajii, K.; Morimoto, H.; *Chem. Pharm. Bull.*, **1988**, *36*, 3281.

165. Lindsay, S., 'High Performance Liquid Chromatography: Analytical Chemistry by Open Learning', Wiley, New York, **1992**.

166. Allen, F. H.; Kennard, O., *Chemical Design Automation News,* **1993**, *8*, 31.